



Università degli Studi di Milano - Bicocca

**DOTTORATO DI RICERCA IN
STATISTICA E APPLICAZIONI**

CICLO XXV

**SORVEGLIANZA DELLA MATRICE DI VARIANZE
E COVARIANZE DI UN PORTAFOGLIO EFFICIENTE
TRAMITE LA TECNICA LASSO**

Settore Scientifico Disciplinare : SECS-S/01

Dottorando

Dott. MARCO GALLOTTI

Tutor

Prof. DIEGO ZAPPA

Anno accademico 2012/2013

INDICE

Introduzione.....	pag. 5
--------------------------	---------------

CAP.1 TEORIA MODERNA DEL PORTAFOGLIO

1.1 Rendimento di un titolo.....	pag. 7
1.2 Variabile casuale R_t	pag. 9
1.3 Portafoglio azionario.....	pag. 10
1.4 Momenti di un portafoglio azionario.....	pag. 12
1.5 Portafoglio con N titoli : diversificazione.....	pag. 19
1.6 Problema di ottimo vincolato.....	pag. 22
1.7 Portafoglio a varianza minima.....	pag. 30
1.8 Portafoglio tangente.....	pag. 32
1.9 Assenza di short sellings.....	pag. 38
1.10 Stime alternative della matrice di varianze e covarianze.....	pag. 40
1.11 Metodi di stima <i>shrinkage</i>	pag. 44

CAP.2 LEAST ABSOLUTE SHRINKAGE AND SELECTION OPERATOR

2.1 Introduzione alla tecnica LASSO.....	pag. 46
2.2 La tecnica LASSO applicata alla matrice di varianze e covarianze.....	pag. 50
2.3 Applicazione.....	pag. 54
2.4 Parametro di penalizzazione ed effetto su pesi e varianza di portafoglio.....	pag. 56

CAP.3 GRAFI

3.1 Cenni alla teoria dei grafi.....	pag. 68
3.2 Applicazione in finanza.....	pag. 71

CAP.4 UTILIZZO DELLA TECNICA LASSO IN FINANZA

4.1 Composizione del portafoglio ottimo.....	pag. 74
4.2 Scelta del parametro di penalizzazione λ	pag. 76
4.3 Dinamiche del mercato azionario.....	pag. 79

CAP.5 SORVEGLIANZA DELLA MATRICE DI VARIANZE E COVARIANZE

5.1 Financial Surveillance.....	pag. 83
5.2 Sorveglianza della matrice di varianze e covarianze dei rendimenti dei titoli di un portafoglio azionario tramite la tecnica LASSO.....	pag. 86
5.3 Costruzione di una carta di controllo : un'applicazione.....	pag. 91

Conclusioni e sviluppi futuri.....	pag. 93
---	----------------

Appendice – Script in R.....	pag. 96
-------------------------------------	----------------

Bibliografia.....	pag. 100
--------------------------	-----------------

INTRODUZIONE

Nel mercato finanziario, un decisore si trova spesso a operare su più possibilità di investimento e per fare ciò deve utilizzare tutte le informazioni a propria disposizione fino a quell'istante temporale, per poter valutare i *rendimenti futuri* delle attività che sono oggetto della sua analisi, nel rispetto del *rischio* di investimento.

La decisione si svolge in un ambito di incertezza, nel quale è necessario affrontare il problema della scelta *ottimale* di un portafoglio azionario, che ha come scopo principale quello di allocare nel miglior modo possibile le risorse a propria disposizione, massimizzando una funzione di utilità che comprenda sia il rendimento atteso (da massimizzare, a parità di rischio) sia il rischio di investimento (da minimizzare, a parità di rendimento) oppure entrambi gli aspetti simultaneamente.

L'idea di questo elaborato nasce dal concetto di “sorveglianza statistica”, trattato nella recente letteratura finanziaria in particolare con gli studi di Frisen (Financial Surveillance, 2008) che si occupa di monitorare i pesi ω_i di un portafoglio ottimo, con l'applicazione delle tecniche di Statistical Process Control (SPC) in ambito finanziario (Schmid, 2010).

In questo contesto, ci si occuperà di monitorare la matrice di varianze e covarianze, Σ , dei rendimenti di titoli appartenenti a un portafoglio azionario, introducendo in particolare una tecnica di stima basata sul metodo LASSO (Least Absolute Shrinkage and Selection Operator).

Diversi autori hanno utilizzato tecniche di stima penalizzata, come la LASSO, applicata ai modelli di regressione, al fine di ridurre il rischio di utilizzare troppi parametri per catturare la dipendenza tra le variabili, ma è soprattutto con i lavori di Tibshirani (2008) che l'algoritmo LASSO viene applicato direttamente alla matrice di varianze e covarianze campionaria, con l'intenzione di aumentare la sua sparsità. Ciò è noto come Graphical Lasso (GLASSO).

Obiettivo di questo lavoro sarà dunque principalmente quello di applicare una metodologia alternativa alla stima della matrice di varianze e covarianze tra i rendimenti dei titoli componenti il portafoglio, al fine di monitorare la composizione del portafoglio stesso per supportare il processo decisionale dell'investitore e mettere in luce le dipendenze tra i titoli.

I risultati che ne conseguiranno mostrano come l'utilizzo della tecnica LASSO in questo contesto permetta di effettuare sia un'operazione di *shrinkage* che di *variable selection* tra i titoli, che consente, oltre ad una migliore allocazione del portafoglio stesso, anche una maggior comprensione del mercato azionario nel suo complesso, in un'ottica spazio-temporale.

Nella prima parte dell'elaborato, dal paragrafo 1.1 al paragrafo 1.9, verranno introdotti alcuni concetti che stanno alla base della teoria moderna del portafoglio (Markowitz, "Portfolio selection", 1952) al fine di fare chiarezza sulle nozioni che verranno utilizzate nel seguito.

Nei due paragrafi a conclusione del primo capitolo verranno messi in luce i problemi derivanti dalla stima della matrice di varianze e covarianze effettuata tramite il campione, evidenziando come numerosi autori nella letteratura moderna abbiano proposto diverse tecniche per la stima della matrice Σ , ponendo l'attenzione soprattutto sulle tecniche di stima *shrinkage*.

La tecnica LASSO verrà quindi presentata nel capitolo 2, prima come tecnica di stima penalizzata nell'ambito della regressione lineare, poi come metodo di stima applicata al contesto finanziario, in cui la penalizzazione viene effettuata direttamente alla matrice di varianze e covarianze Σ .

Gli effetti che tale procedura implica sull'ottimizzazione di portafoglio verranno mostrati nei successivi due capitoli. Anzitutto, nel capitolo 3, si utilizzerà uno strumento derivante dalla teoria dei grafi che permetterà di comprendere maggiormente le dipendenze tra i titoli azionari e, successivamente, nel capitolo 4, si mostrerà tramite esempi applicativi come l'utilizzo della tecnica LASSO possa apportare benefici in termini di comprensione e ottimizzazione di portafoglio.

Nel quinto ed ultimo capitolo, l'obiettivo sarà quello di sorvegliare la matrice di varianze e covarianze tramite carte di controllo multivariate, derivanti dalle tecniche di *Financial Surveillance*, con un'applicazione al contesto finanziario.

Infine verranno mostrati, all'interno delle conclusioni, alcuni sviluppi futuri che potranno essere affrontati alla luce dei risultati presentati nell'elaborato.

CAP. 1

TEORIA MODERNA DEL PORTAFOGLIO

1.1 Rendimento di un titolo

Si considerino N titoli azionari o obbligazionari.

Si definisce $R_{i,t}$ il tasso di rendimento al tempo t dell'i-esimo titolo, per $i=1,2,\dots,N$, come :

$$R_{i,t} = \frac{P_{i,t} - P_{i,t-1} + D_{i,t}}{P_{i,t-1}}$$

dove:

$P_{i,t} > 0$ è il prezzo (quotazione) del titolo all'istante t;

$P_{i,t-1} > 0$ è il prezzo (quotazione) del titolo all'istante t-1;

$D_{i,t} > 0$ è la cedola (o dividendo) ed eventuali premi previsti all'istante t.

Per semplificare l'analisi non verranno presi in considerazione, nel seguito, premi e dividendi. Gli intervalli temporali saranno ravvicinati nel tempo, tipicamente a cadenza giornaliera.

Quindi :

$$R_{i,t} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} = \frac{P_{i,t}}{P_{i,t-1}} - 1$$

da cui si ha $R_{i,t} > -1$.

In letteratura finanziaria, si fa un ampio uso del log-rendimento di un titolo, ponendo :

$$\text{Ln} \left(\frac{P_{i,t}}{P_{i,t-1}} \right) = \text{Ln} (R_{i,t} + 1)$$

D'ora in avanti, nel corso della trattazione, si parlerà di “rendimenti” sottintendendo il riferimento ai log-rendimenti di un titolo azionario.

Il passaggio dal rendimento al log-rendimento è necessario per più di un motivo.

Anzitutto, l'osservazione empirica e alcune considerazioni teoriche¹ suggeriscono che solitamente i prezzi abbiano una distribuzione fortemente asimmetrica e sempre positiva, identificabile con una variabile casuale log-normale.

Sotto questa assunzione, è noto che il logaritmo di una distribuzione log-normale si distribuisce come una normale, quindi il rendimento può essere visto come :

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln(P_t) - \ln(P_{t-1})$$

ovvero una combinazione lineare di variabili casuali normali che, come è noto, è ancora una variabile casuale normale. Di qui l'assunto di normalità distributiva dei rendimenti.

Inoltre, in assenza di ipotesi parametriche, è noto che approssimando la funzione $\ln(x)$ in un intorno di $x = 1$ tramite lo sviluppo in serie di Taylor troncata al primo termine si ottiene $\ln(1 + x) \approx x$. Ponendo dunque $x = R_t = \frac{P_t}{P_{t-1}} - 1$ si ha che $\ln(R_t + 1)$ è in generale una buona approssimazione del rendimento stesso R_t , ovvero :

$$\ln\left(\frac{P_t}{P_{t-1}}\right) \approx \frac{P_t}{P_{t-1}} - 1$$

Il passaggio al log-rendimento consente inoltre di calcolare il rendimento di un titolo azionario al tempo t rispetto al tempo $(t-k)$, facendo dunque riferimento a un intervallo temporale di ampiezza $k > 1$, come somma dei singoli rendimenti :

$$R_{t,(t-k)} = \ln\left(\frac{P_t}{P_{t-k}}\right) = \ln\left(\frac{P_t}{P_{t-1}} \cdot \frac{P_{t-1}}{P_{t-2}} \cdot \dots \cdot \frac{P_{t-k+1}}{P_{t-k}}\right) = \sum_{i=1}^k \frac{P_{t-i+1}}{P_{t-i}} = \sum_{i=1}^k R_{(t-i+1),(t-i)}$$

¹ Si veda "Options, futures and other derivatives, 8th edition, cap.13-14" John C. Hull, 2012.

1.2 Variabile casuale R_t

Il rendimento di un titolo azionario al tempo $t+1$, R_{t+1} , non è certo, ma può essere visto come una *variabile casuale*.

Quale sia il tipo di distribuzione di probabilità del rendimento di un titolo in generale non è noto, ma alcune verifiche empiriche hanno portato a ricorrere spesso alla distribuzione Gaussiana, anche per le motivazioni sopracitate.

L'approccio classico considera quindi R_t come una variabile casuale normale, caratterizzata dal valor medio μ_r , che misura il rendimento atteso sul titolo, e dallo scarto quadratico medio σ_r , che ne misura la dispersione.

Lo scarto quadratico medio σ_r fu proposto già da Markowitz nel 1952 come misura del rischio d'investimento; infatti è noto che la redditività di un'attività finanziaria si considera tanto più incerta quanto più è elevata la probabilità che i rendimenti futuri si disperdano rispetto al valore atteso.

Il *rischio* di una singola attività finanziaria si calcola dunque come :

$$s_r^2 = \frac{1}{T-1} \sum_{t=1}^T (R_t - \bar{R})^2$$

dove il rendimento atteso del titolo può essere stimato tramite :

$$\bar{R} = \frac{1}{T} \sum_{t=1}^T R_t$$

Si noti che la letteratura recente sta rivolgendo sempre più spesso l'attenzione all'allocazione ottimale di portafoglio *a quattro dimensioni*, ovvero che tenga in considerazione i primi quattro momenti centrali, analizzando anche asimmetria e curtosi dei rendimenti. Tali aspetti non vengono tuttavia considerati nel presente elaborato, il cui obiettivo principale è quello di focalizzare l'attenzione sulla matrice di varianze e covarianze tra i titoli presenti in portafoglio.

1.3 Portafoglio azionario

In finanza, un *portafoglio* è genericamente definito come un insieme di strumenti (conti correnti, azioni, obbligazioni, fondi comuni, ecc.) a disposizione di un investitore.

Da questo momento in avanti, onde evitare complicazioni nell'interpretazione delle caratteristiche monetarie sottostanti, ci si occuperà esclusivamente di portafogli *azionari*, ovvero portafogli composti esclusivamente da titoli azionari.

Ogni singolo titolo che compone il portafoglio, avrà dunque al tempo t un rendimento atteso pari a $\bar{R}$ ed una misura di rischio data da s_r .

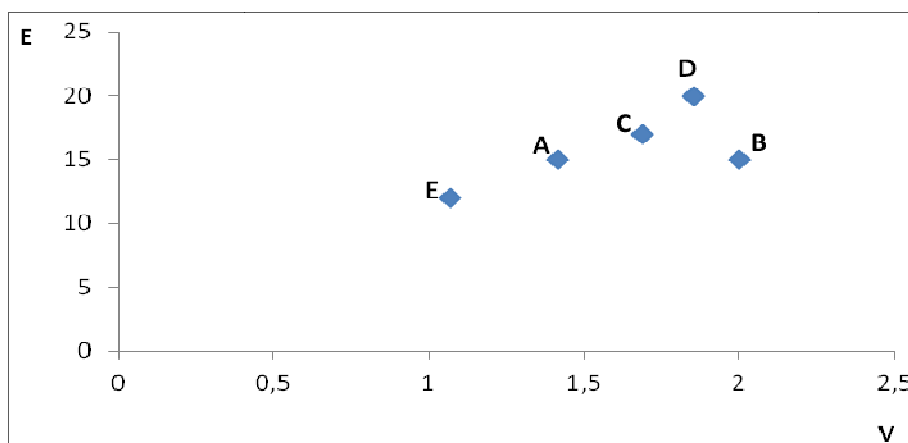
Obiettivo principale di qualsiasi decisore è quello di *allocare* nel miglior modo possibile le risorse a propria disposizione, per investire maggiormente su titoli a rendimento atteso più elevato e rischio minore, a discapito ovviamente di titoli più rischiosi e con rendimento atteso inferiore.

Un primo passo per raggiungere tale scopo può essere quello di posizionare i singoli titoli all'interno di un grafico rischio-rendimento (noto in letteratura come grafico "E-V", ovvero Expected value-Variance) in cui sull'asse delle ordinate vengano riportati i rendimenti attesi [$E = \bar{R}$] e sull'asse delle ascisse lo scarto quadratico medio di ciascun titolo [$V = s_r$].

Esempio 1.

Si considerino 5 titoli azionari, i cui rendimenti vengono riassunti nel seguente prospetto :

	A	B	C	D	E
E	15	15	17	20	12
V	1,41	2	1,69	1,85	1,07



Si può notare che più un titolo è posizionato nella regione *Nord-Ovest* del quadrante, e più è preferibile dall'investitore, perché assicura un maggior valore atteso del rendimento, congiuntamente ad un minor rischio.

Il principio della dominanza², suggerito proprio a partire dal grafico E-V, permette di ordinare (ovvero preferire) i titoli che compongono il portafoglio nel seguente modo :

- 1) a parità di rendimento atteso, viene scelta l'opportunità meno rischiosa, ossia con minor scarto quadratico medio (dominanza per rischio);
- 2) a parità di scarto quadratico medio, viene scelta l'opportunità con rendimento atteso maggiore (dominanza per rendimento);
- 3) nel caso in cui $\bar{R}_i < \bar{R}_j$ ed $s_i < s_j$ il titolo j-esimo non domina il titolo i-esimo.

Si prenda ora in considerazione l'Esempio 1. In questo caso, il principio della dominanza non aiuta il decisore nella scelta di un titolo, poiché non attua un ordinamento totale, ma solamente parziale, mantenendo indecisione. L'applicazione del criterio della dominanza alle cinque opportunità dell'Esempio 1 porta infatti a verificare che :

- 1) nessun decisore sceglierà B, perché dominato da A per rischio e da C e D sia per rendimento che per rischio;
- 2) E, A, C, D sono non dominati da nessun altro titolo, poiché a maggior rendimento corrisponde maggior rischio, e come si vede in figura non esiste alcun titolo posto nella regione più a Nord-Ovest di loro stessi.

Scartato B, non esiste quindi, tra A, C, D, E, una soluzione ritenuta migliore delle altre.

Quando si ha a che fare con un generico portafoglio azionario, composto da più titoli (generalmente *correlati* tra loro) un investitore si troverà spesso a dover far fronte ad una situazione simile a quella descritta nell'esempio, quindi sarà costretto a ripartire le risorse a propria disposizione scegliendo opportuni pesi ω_i per ogni singolo titolo i , al fine di ottenere il massimo rendimento atteso del portafoglio a parità di rischio, oppure il minor rischio a parità di rendimento atteso.

² cfr. Markowitz (1987)

1.4 Momenti di un portafoglio azionario

Un portafoglio azionario P , composto da N titoli, è caratterizzato principalmente da due indicatori : media e varianza.

Il rendimento atteso del portafoglio, μ_P , è dato dalla media aritmetica ponderata dei rendimenti attesi dei singoli titoli che lo compongono :

$$E(R_P) = \mu_P = \sum_{i=1}^N \omega_i \mu_i$$

dove :

ω_i è la quantità di risorse economiche investita nell' i -esimo titolo.

La varianza di portafoglio, σ_P^2 , è data dal valore atteso del quadrato degli scarti del rendimento di portafoglio dal rendimento atteso del portafoglio stesso, ovvero :

$$Var(R_P) = \sigma_P^2 = \sum_{i=1}^N \omega_i \sigma_i^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \omega_i \omega_j \sigma_i \sigma_j \rho_{ij}$$

dove :

ρ_{ij} è il coefficiente di correlazione tra il titolo i -esimo e quello j -esimo.

Al fine di ottenere la migliore allocazione possibile delle risorse a propria disposizione, il problema principale è dunque la stima ottimale dei pesi ω_i sotto opportune condizioni :

1. La somma dei pesi dev'essere unitaria, poiché ogni ω_i rappresenta la percentuale di ricchezza da allocare al titolo i -esimo. Quindi $\sum_{i=1}^N \omega_i = 1$.
2. se si esclude la presenza di *short sellings*³, ogni singolo peso deve necessariamente essere positivo. Quindi $\omega_i \geq 0 \quad \forall i$.

³ Per short selling nella letteratura finanziaria si intende la vendita allo scoperto (effettuata nei confronti di soggetti terzi) di titoli non direttamente posseduti dal venditore. Effettuando questo tipo di operazione, l'intento è quello di ottenere un profitto a seguito di un movimento ribassista delle quotazioni, che permetta di acquisire il titolo non posseduto ad un prezzo vantaggioso rispetto al momento della vendita allo scoperto.

Si può notare che il rischio (da minimizzare) associato a un portafoglio con più titoli, sia dipendente non solo dal rischio intrinseco di ogni singolo titolo, ma anche dalle correlazioni lineari esistenti tra di essi. Ecco perché è necessario stimare e monitorare la matrice di varianze e covarianze Σ tra i titoli che compongono il portafoglio, che sarà oggetto del presente elaborato.

A titolo esemplificativo, si parta inizialmente da un portafoglio composto da soli 2 titoli azionari.

In tal caso, il rendimento del portafoglio è pari a :

$$R_P = \omega_1 R_1 + \omega_2 R_2$$

La varianza del portafoglio risulta :

$$\sigma_P^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2\omega_1 \omega_2 Cov[R_1, R_2]$$

ma essendo : $\rho = \frac{Cov[R_1, R_2]}{\sigma_1 \sigma_2}$

si ottiene : $\sigma_P^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2\omega_1 \omega_2 \rho \sigma_1 \sigma_2$

che lega il rischio del portafoglio all'indice di correlazione lineare ρ tra i due titoli.

A seconda del particolare valore del coefficiente di correlazione ρ , si possono distinguere diversi casi⁴, nei quali il rischio e il rendimento del portafoglio danno origine a particolari combinazioni.

CASO 1 : Correlazione perfettamente positiva ($\rho = 1$) :

$$\sigma_P^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2\omega_1 \omega_2 \sigma_1 \sigma_2 = (\omega_1 \sigma_1 + \omega_2 \sigma_2)^2$$

da cui : $\sigma_P = \omega_1 \sigma_1 + \omega_2 \sigma_2$

In questo caso la misura del rischio di portafoglio è data da una combinazione lineare dei singoli rischi, così come il rendimento è una combinazione lineare dei singoli rendimenti :

⁴ cfr. Elton and Gruber (1981)

$$R_P = \omega_1 R_1 + \omega_2 R_2$$

Mettendo a sistema le due equazioni e ricordando che $\omega_1 = (1 - \omega_2)$ si può ottenere ω_1 in funzione del rischio :

$$\begin{aligned}\sigma_P &= \omega_1 \cdot \sigma_1 + \omega_2 \cdot \sigma_2 = \omega_1 \cdot \sigma_1 + (1 - \omega_1) \cdot \sigma_2 = \\ &= \omega_1 \cdot \sigma_1 + \sigma_2 - \omega_1 \cdot \sigma_2 = \omega_1(\sigma_1 - \sigma_2) + \sigma_2\end{aligned}$$

che porta a :

$$\omega_1 = \frac{\sigma_P - \sigma_2}{\sigma_1 - \sigma_2}$$

Quindi il rendimento atteso del portafoglio, in funzione del rischio, è dato da :

$$\begin{aligned}\mu_P &= \omega_1 \cdot \mu_1 + (1 - \omega_1) \cdot \mu_2 = \\ &= \left(\frac{\sigma_P - \sigma_2}{\sigma_1 - \sigma_2} \right) \cdot \mu_1 + \left(1 - \frac{\sigma_P - \sigma_2}{\sigma_1 - \sigma_2} \right) \cdot \mu_2 = \\ &= \frac{\sigma_P \mu_1 - \sigma_2 \mu_1}{\sigma_1 - \sigma_2} + \mu_2 - \frac{\sigma_P \mu_2 - \sigma_2 \mu_2}{\sigma_1 - \sigma_2} = \\ &= \mu_2 + \frac{\sigma_P \mu_1 - \sigma_2 \mu_1 - \sigma_P \mu_2 + \sigma_2 \mu_2}{\sigma_1 - \sigma_2} = \\ &= \left(\mu_2 - \frac{\mu_1 - \mu_2}{\sigma_1 - \sigma_2} \sigma_2 \right) + \left(\frac{\mu_1 - \mu_2}{\sigma_1 - \sigma_2} \right) \sigma_P\end{aligned}$$

Ciò significa che dato il rendimento atteso e il rischio di ciascun titolo, il rendimento atteso del portafoglio è linearmente dipendente dal suo rischio. Dunque, all'aumentare del rischio aumenta il rendimento atteso del portafoglio e viceversa.

Essendo inoltre R_P e σ_P interpretabili come medie ponderate rispettivamente dei singoli R_i e dei singoli σ_i , non è possibile trovare un portafoglio con rischio minore del rischio minimo tra i due titoli, né un portafoglio con rendimento atteso maggiore del rendimento atteso massimo tra i due, il che rende questa situazione poco favorevole per l'investitore.

CASO 2 : Correlazione perfettamente negativa ($\rho = -1$) :

$$\sigma_P^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 - 2\omega_1 \omega_2 \sigma_1 \sigma_2 = (\omega_1 \sigma_1 - \omega_2 \sigma_2)^2$$

da cui, ponendo $\omega_1 \sigma_1 > \omega_2 \sigma_2$ si ha :

$$\begin{aligned}\sigma_P &= \omega_1 \cdot \sigma_1 - \omega_2 \cdot \sigma_2 = \omega_1 \cdot \sigma_1 - (1 - \omega_1) \cdot \sigma_2 = \\ &= \omega_1 \cdot \sigma_1 - \sigma_2 + \omega_1 \cdot \sigma_2 = \omega_1 (\sigma_1 + \sigma_2) - \sigma_2\end{aligned}$$

$$\omega_1 = \frac{\sigma_P + \sigma_2}{\sigma_1 + \sigma_2}$$

che porta ad un rendimento atteso del portafoglio pari a :

$$\begin{aligned}\mu_P &= \omega_1 \cdot \mu_1 + (1 - \omega_1) \cdot \mu_2 = \\ &= \omega_1 \cdot \mu_1 + \mu_2 - \omega_1 \mu_2 = \\ &= \mu_2 + (\mu_1 - \mu_2) \cdot \omega_1 = \\ &= \mu_2 + (\mu_1 - \mu_2) \cdot \left(\frac{\sigma_P + \sigma_2}{\sigma_1 + \sigma_2} \right)\end{aligned}$$

Mentre nel secondo caso, ovvero con $\omega_1 \sigma_1 < \omega_2 \sigma_2$, si ha :

$$\begin{aligned}\sigma_P &= \omega_2 \cdot \sigma_2 - \omega_1 \cdot \sigma_1 = (1 - \omega_1) \cdot \sigma_2 - \omega_1 \cdot \sigma_1 = \\ &= \sigma_2 - \omega_1 \cdot \sigma_2 - \omega_1 \cdot \sigma_1 = \sigma_2 - \omega_1 (\sigma_1 + \sigma_2)\end{aligned}$$

$$\omega_1 = \frac{\sigma_2 - \sigma_P}{\sigma_1 + \sigma_2}$$

$$\mu_P = \omega_1 \cdot \mu_1 + (1 - \omega_1) \cdot \mu_2 = \mu_2 + (\mu_1 - \mu_2) \cdot \left(\frac{\sigma_2 - \sigma_P}{\sigma_1 + \sigma_2} \right)$$

Quindi, nel caso di perfetta correlazione negativa, esiste un portafoglio a rischio nullo, cioè con varianza pari a 0, che si ottiene ponendo :

$$\omega_1^* = \frac{\sigma_2}{\sigma_1 + \sigma_2} \quad \omega_2^* = \frac{\sigma_1}{\sigma_1 + \sigma_2}$$

che porta ad un rendimento atteso pari a :

$$\mu_P = \mu_2 + (\mu_1 - \mu_2) \cdot \left(\frac{\sigma_2}{\sigma_1 + \sigma_2} \right)$$

CASO 3 : Correlazione nulla ($\rho = 0$) :

$$\sigma_P^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2$$

Essendo questa espressione una combinazione lineare delle varianze dei singoli titoli, non può essere annullata, ma può essere facilmente minimizzata, ponendo uguale a zero la seguente espressione :

$$\frac{\partial \sigma^2}{\partial \omega_2} = -2(1 - \omega_2)\sigma_1^2 + 2\omega_2\sigma_2^2 = 0$$

che equivale a :

$$\omega_2\sigma_1^2 + \omega_2\sigma_2^2 - \sigma_1^2 = 0$$

da cui :

$$\omega_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \quad e \quad \omega_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

In questo caso la varianza del portafoglio diventa pari a :

$$\sigma_P^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

da cui consegue che : $\sigma_P^2 < \sigma_1^2$ e $\sigma_P^2 < \sigma_2^2$

Mentre il rendimento atteso risulta essere :

$$\begin{aligned} \mu_P &= \omega_1 \cdot \mu_1 + (1 - \omega_1) \cdot \mu_2 = \\ &= \left(\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \right) \cdot \mu_1 + \left(1 - \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \right) \cdot \mu_2 \end{aligned}$$

CASO 4 : Correlazione qualsiasi ($-1 < \rho < 1$, ma $\rho \neq 0$) :

$$\sigma_P^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2\omega_1 \omega_2 \rho \sigma_1 \sigma_2$$

Si nota che, date le varianze dei singoli titoli, il coefficiente di correlazione lineare tra i titoli, ρ , gioca un ruolo chiave nel calcolo del rischio, poiché più ρ assume un valore elevato, maggiore sarà il rischio del portafoglio.

Si torni ora al caso generale con più di 2 titoli azionari, ovvero l'Esempio 1, in cui $N = 5$.

La matrice Σ di varianze e covarianze tra i 5 titoli si assuma essere la seguente :

	A	B	C	D	E
A	2	-0,4286	2,2857	0,1429	-0,2857
B	-0,4286	4	-0,2857	1,1429	1,1429
C	2,2857	-0,2857	2,8571	0,7143	-0,5714
D	0,1429	1,1429	0,7143	3,4286	0,1429
E	-0,2857	1,1429	-0,5714	0,1429	1,1429

Ipotizzando di allocare in modo uniforme le risorse a propria disposizione, ovvero di porre il 20% del capitale in ciascuna attività, è possibile calcolare il valore atteso e la varianza del portafoglio :

$$\mu_P = \sum_{i=1}^N \omega_i \cdot \mu_i = 0,2 \cdot \sum_{i=1}^N \mu_i = \mathbf{15,8}$$

$$\sigma_P^2 = \sum_{i=1}^N \omega_i^2 \cdot \sigma_i^2 + 2 \cdot \sum_{i=1}^N \sum_{j=i+1}^N \omega_i \omega_j \cdot \sigma_{ij} = 0,537 + 2 \cdot 0,12 = \mathbf{0,777}$$

Una prima analisi, tramite il grafico E-V, aveva permesso di effettuare alcune conclusioni preliminari riguardanti i singoli titoli A,B,C,D,E, che potrebbe far propendere in prima istanza ad una *diversa allocazione* delle risorse tra i titoli, in modo tale da aumentare il rendimento atteso del portafoglio e/o di minimizzarne il rischio.

Se per esempio, si ponessero le risorse a propria disposizione in modo tale da allocare maggior peso ad A e E (poiché meno rischiosi) a discapito di B, C, D, per esempio :

35% in A 10% in B 10% in C 10% in D 35% in E

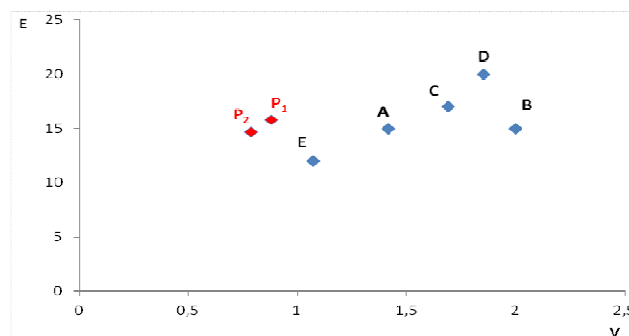
si otterrebbe il seguente valore atteso del portafoglio :

$$\mu_P = \sum_{i=1}^N \omega_i \cdot \mu_i = 0,35 \cdot 15 + 0,1 \cdot 15 + 0,1 \cdot 17 + 0,1 \cdot 20 + 0,35 \cdot 12 = \mathbf{14,65}$$

e la seguente misura di rischio :

$$\sigma_P^2 = \sum_{i=1}^N \omega_i^2 \cdot \sigma_i^2 + 2 \cdot \sum_{i=1}^N \sum_{j=i+1}^N \omega_i \omega_j \cdot \sigma_{ij} = 0,4878 + 2 \cdot 0,0657 = \mathbf{0,619}$$

Questa seconda allocazione, indotta da un semplice ragionamento grafico, sembra essere migliore della prima in termini di rischio, in quanto, pur misurando un minor rendimento atteso, produce un rischio di portafoglio minore, come si può visualizzare nel grafico E-V :



dove P₁ è il portafoglio con ripartizione uniforme del capitale per ogni titolo, mentre P₂ rappresenta il portafoglio con la seconda allocazione, che risulta essere meno rischiosa.

Questo esempio introduttivo aiuta il lettore a comprendere come l'obiettivo sia quello di stimare il vettore di pesi ω_i ottimale, per trovare il portafoglio meno rischioso, oppure con rendimento atteso più elevato, oppure per soddisfare entrambe le condizioni simultaneamente : minimizzare il rischio fissando un rendimento atteso oppure massimizzare il rendimento atteso fissando un livello di rischio.

Come abbiamo potuto notare, giocano un ruolo fondamentale in quest'ottica le covarianze tra i titoli, quindi la stima della matrice di varianze e covarianze sarà centrale nella trattazione.

1.5 Portafoglio con N titoli : diversificazione

I risultati ottenuti nel precedente paragrafo mostrano come sia possibile ottenere un portafoglio con rischio σ_P minore rispetto a tutti i rischi dei singoli titoli che lo compongono, tranne nel caso particolare di perfetta correlazione positiva tra i titoli (ovvero $\rho=1$).

Un'allocazione ottimale delle risorse permette quindi all'investitore di rischiare meno investendo su più titoli contemporaneamente, piuttosto che su uno qualsiasi di essi.

Il ragionamento che sta alla base della diversificazione è dato dalla sub-additività dei rischi di investimento, la quale afferma che⁵ :

$$\sqrt{\text{Var}(R_1 + R_2)} \leq \sqrt{\text{Var}(R_1)} + \sqrt{\text{Var}(R_2)}$$

Ovvero, presi due rendimenti R_1 ed R_2 , lo scarto quadratico medio dell'intero portafoglio (R_1+R_2) non è mai più elevato della somma degli scarti quadratici medi dei singoli rendimenti che compongono il portafoglio.

In altre parole, la diversificazione degli investimenti riduce il rischio del portafoglio : aggiungendo titoli a un portafoglio, l'indice di rischio dell'intero portafoglio non aumenta più della somma degli indici di rischio dei nuovi titoli inseriti.

Tale proprietà si dimostra facilmente, poiché elevando entrambi i membri al quadrato si ottiene :

$$\text{Var}(R_1 + R_2) \leq \text{Var}(R_1) + \text{Var}(R_2) + 2\sqrt{\text{Var}(R_1)}\sqrt{\text{Var}(R_2)}$$

Ovvero :

$$\begin{aligned} \text{Var}(R_1) + \text{Var}(R_2) + 2 \text{Cov}(R_1, R_2) \\ \leq \text{Var}(R_1) + \text{Var}(R_2) + 2\sqrt{\text{Var}(R_1)}\sqrt{\text{Var}(R_2)} \end{aligned}$$

Semplificando :

$$\text{Cov}(R_1, R_2) \leq \sqrt{\text{Var}(R_1)}\sqrt{\text{Var}(R_2)}$$

che è sempre valida per la disuguaglianza di Cauchy-Schwarz.

⁵ cfr. "Coherent measures of risk", Artzner et al (1998)

Dividendo quindi l'investimento su più titoli si riduce il rischio, soprattutto se, come abbiamo già visto, essi sono incorrelati tra loro, o ancor meglio, correlati negativamente, poiché, quando un titolo mostra difficoltà, gli altri titoli in portafoglio compensano le perdite, garantendo un rendimento medio complessivo più alto.

Il risultato finale della diversificazione risulta essere la media dei rendimenti migliori e peggiori, diminuendo quindi il rischio di perdite.

Si prenda ad esempio il caso limite in cui i titoli siano incorrelati tra loro.

Si ha che :

$$\sigma_P^2 = \sum_i^N \omega_i^2 \cdot \sigma_i^2$$

Supponendo di equiripartire l'investimento tra gli N titoli, cioè $\omega_i = \frac{1}{N}$ per ogni i, la varianza di portafoglio risulta :

$$\sigma_P^2 = \sum_{i=1}^N \left(\frac{1}{N}\right)^2 \sigma_i^2 = \frac{1}{N} \left[\sum_{i=1}^N \frac{\sigma_i^2}{N} \right]$$

Il termine entro parentesi quadra è semplicemente la varianza media dei titoli, ovvero la cosiddetta varianza “within” (media delle varianze, simboleggiata con σ_w^2) ed è facile verificare che al divergere di N (quindi all'aumentare del numero di titoli compresi nel portafoglio) essa tende a zero, così come, di conseguenza, la varianza del portafoglio :

$$\lim_{N \rightarrow \infty} \sigma_P^2 = \lim_{N \rightarrow \infty} (1/N) \sigma_w^2 = 0$$

Anche nel caso di titoli perfettamente correlati negativamente ($\rho=-1$) abbiamo già visto con un semplice esempio (N=2) come la varianza totale del portafoglio possa essere completamente annullata.

Quindi, operando con infiniti titoli tutti incorrelati linearmente nei rendimenti, oppure con titoli esattamente correlati negativamente, si riesce ad annullare il rischio di portafoglio, ma nella realtà è poco probabile investire su infiniti titoli che risultino essere tutti incorrelati tra loro oppure titoli esattamente correlati negativamente.

Nel caso generale ($p \neq 0$), infatti, si ha che, ipotizzando sempre un'equiripartizione della ricchezza tra gli N titoli, la varianza del portafoglio è pari a :

$$\sigma_P^2 = \sum_{i=1}^N \left(\frac{1}{N}\right)^2 \sigma_i^2 + \sum_{i=1}^N \sum_{j \neq i}^N \left(\frac{1}{N}\right) \left(\frac{1}{N}\right) \sigma_{ij}$$

da cui :

$$\sigma_P^2 = \frac{1}{N} \left[\sum_{i=1}^N \frac{\sigma_i^2}{N} \right] + \frac{N-1}{N} \sum_{i=1}^N \sum_{j \neq i}^N \left[\frac{\sigma_{ij}}{N(N-1)} \right]$$

Sostituendo i termini tra parentesi, l'espressione può essere riscritta come :

$$\sigma_P^2 = \frac{1}{N} \sigma_w^2 + \frac{N-1}{N} \overline{\sigma_{ij}}$$

dove $\overline{\sigma_{ij}}$ rappresenta la *covarianza media* dei titoli in portafoglio, essendo le covarianze in numero di $N(N-1)$.

Allora per $N \rightarrow \infty$, si ha che il primo addendo tende a zero, mentre il secondo addendo tende al valore $\overline{\sigma_{ij}}$ che non può essere annullato ed è infatti noto come "rischio sistematico"⁶ :

$$\lim_{N \rightarrow \infty} \sigma_P^2 = \lim_{N \rightarrow \infty} \left[\frac{1}{N} \sigma_w^2 + \frac{N-1}{N} \overline{\sigma_{ij}} \right] = \overline{\sigma_{ij}}$$

⁶ cfr. Elton and Gruber (1984)

1.6 Problema di ottimo vincolato

Fino a questo momento la media e la varianza di un portafoglio azionario composto da N titoli sono state presentate come :

$$\mu_P = \sum_{i=1}^N \omega_i \mu_i$$

$$\sigma_P^2 = \sum_{i=1}^N \omega_i^2 \sigma_i^2 + 2 \sum_{i=1}^N \sum_{j>i}^N \omega_i \omega_j \sigma_{ij}$$

Ponendo la varianza di ogni singolo titolo $\text{Var}(R_i) = \sigma_i^2 = \sigma_{ii}$, l'espressione della varianza di portafoglio può essere semplificata in :

$$\sigma_P^2 = \sum_{i=1}^N \sum_{j=1}^N \omega_i \omega_j \sigma_{ij}$$

Ponendo ora :

$$\boldsymbol{\omega} = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \dots \\ \omega_N \end{bmatrix}$$

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \dots \\ \mu_N \end{bmatrix}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \dots & \sigma_{NN} \end{bmatrix}$$

le espressioni di media e varianza possono essere facilmente rappresentate in nozione matriciale.

Infatti :

$$\mu_P = \boldsymbol{\omega}' \boldsymbol{\mu}$$

$$\sigma_P^2 = \boldsymbol{\omega}' \boldsymbol{\Sigma} \boldsymbol{\omega}$$

DEFINIZIONE :

Un portafoglio azionario si definisce *efficiente* quando è caratterizzato dalla varianza più bassa tra tutti i portafogli che hanno lo stesso rendimento atteso.

Dunque, per un dato rendimento atteso μ_P , il portafoglio efficiente è dato dal vettore di pesi

$\omega' = [\omega_1, \omega_2, \dots, \omega_N]$ che risolve il seguente problema di ottimo vincolato :

$$\begin{aligned} \min \quad & \sigma_P^2 = \sum_i \sum_j \omega_i \omega_j \sigma_{ij} \\ \text{sub} \quad & \mu_P = \sum_i \omega_i \mu_i \\ & \sum_i \omega_i = 1 \end{aligned}$$

In forma matriciale :

$$\begin{aligned} \min \quad & \omega' \Sigma \omega \\ \text{sub} \quad & \omega' \mu = \mu_P \\ & \omega' \bar{\mathbf{1}} = 1 \end{aligned}$$

Dove con $\bar{\mathbf{1}}$ si indica un vettore (Nx1) composto di 1 : $\bar{\mathbf{1}} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$

La minimizzazione vincolata si risolve utilizzando il Lagrangiano :

$$L(\omega, \delta_1, \delta_2) = \omega' \Sigma \omega + \delta_1 (\mu_P - \omega' \mu) + \delta_2 (1 - \omega' \bar{\mathbf{1}})$$

dove δ_1 e δ_2 sono i moltiplicatori di Lagrange e ω_P è il vettore di pesi del portafoglio efficiente.

Quindi si risolve :

$$\frac{\partial L(\omega, \delta_1, \delta_2)}{\partial \omega} = 2\Sigma\omega_P - \delta_1\mu - \delta_2\bar{\mathbf{1}} = 0$$

Ovvero :

$$\omega_P = \frac{1}{2} \Sigma^{-1} (\delta_1 \mu + \delta_2 \bar{\mathbf{1}}) = \Sigma^{-1} (\lambda_1 \mu + \lambda_2 \bar{\mathbf{1}}) = \lambda_1 \Sigma^{-1} \mu + \lambda_2 \Sigma^{-1} \bar{\mathbf{1}}$$

dove $\lambda_1 = \delta_1/2$ e $\lambda_2 = \delta_2/2$ sono i nuovi moltiplicatori di Lagrange da determinare utilizzando i vincoli della minimizzazione :

$$\mu_P = \omega'_P \mu = \mu' \omega_P = \lambda_1 \mu' \Sigma^{-1} \mu + \lambda_2 \mu' \Sigma^{-1} \bar{\mathbf{1}}$$

e

$$1 = \omega'_P \bar{\mathbf{1}} = \bar{\mathbf{1}}' \omega_P = \lambda_1 \bar{\mathbf{1}}' \Sigma^{-1} \mu + \lambda_2 \bar{\mathbf{1}}' \Sigma^{-1} \bar{\mathbf{1}}$$

che rappresenta un sistema di due equazioni in due incognite λ_1 e λ_2 .

Ponendo ora :

$$A = \mu' \Sigma^{-1} \bar{\mathbf{1}} = \bar{\mathbf{1}}' \Sigma^{-1} \mu$$

$$B = \mu' \Sigma^{-1} \mu$$

$$C = \bar{\mathbf{1}}' \Sigma^{-1} \bar{\mathbf{1}}$$

il sistema può essere riscritto come :

$$\mu_P = B \lambda_1 + A \lambda_2$$

$$1 = A \lambda_1 + C \lambda_2$$

e risolto ponendo :

$$\lambda_1 = \frac{-A + C \mu_P}{D}$$

$$\lambda_2 = \frac{B - A \mu_P}{D}$$

dove $D = BC - A^2$ è il determinante della matrice dei coefficienti associata al sistema.

Si ha dunque :

$$\omega_P = \mathbf{g} + \mu_P \mathbf{h}$$

dove :

$$\mathbf{g} = \frac{B \Sigma^{-1} \bar{\mathbf{1}} - A \Sigma^{-1} \mu}{D} = \frac{B}{D} \Sigma^{-1} \bar{\mathbf{1}} - \frac{A}{D} \Sigma^{-1} \mu$$

$$\mathbf{h} = \frac{C \Sigma^{-1} \mu - A \Sigma^{-1} \bar{\mathbf{1}}}{D} = \frac{C}{D} \Sigma^{-1} \mu - \frac{A}{D} \Sigma^{-1} \bar{\mathbf{1}}$$

Si noti che le quantità **g** ed **h** dipendono esclusivamente dal vettore **μ** e soprattutto dalla matrice di varianze e covarianze **Σ**, la cui stima risulta quindi ancora una volta fondamentale e su cui si concentrerà l'attenzione nelle prossime sezioni.

Una volta stimati il vettore **μ** e la matrice **Σ**, si ha che il vettore di pesi efficiente **ω_P** dipende dal valore del rendimento atteso desiderato μ_P .

Per ogni $\mu_{min} \leq \mu_P \leq \mu_{max}$ si ha quindi un vettore di pesi che minimizza la varianza σ_P^2 .

Esempio 2.

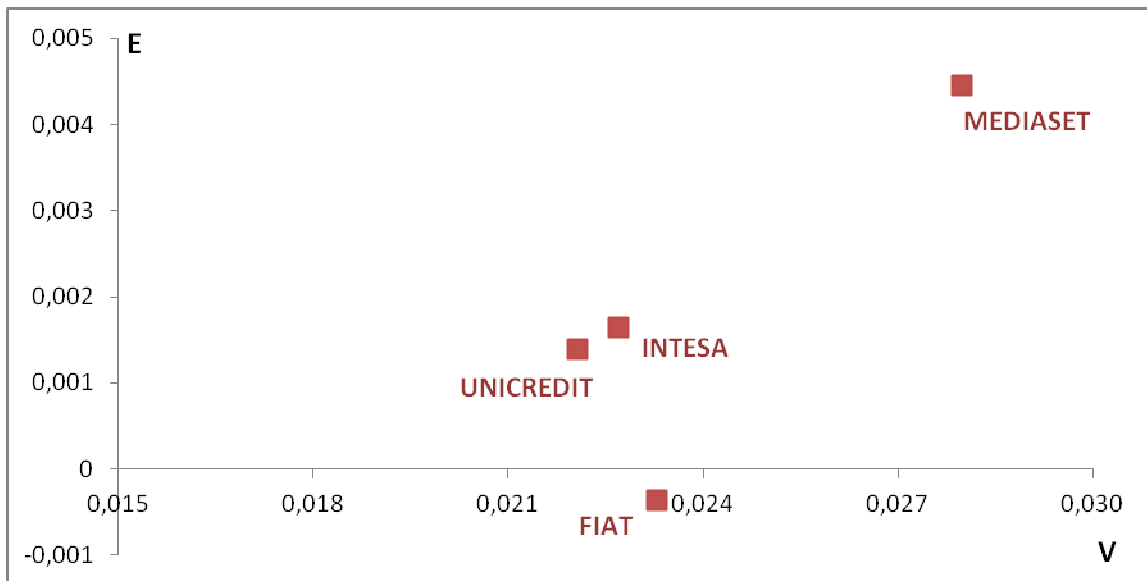
Si prendano i prezzi di chiusura (Adj. close) di N=4 titoli azionari facenti parte dell'indice FTSEMIB della Borsa italiana :

- Intesa San Paolo
- Unicredit
- Fiat
- Mediaset

nel periodo dal 3/6/2013 al 17/9/2013 per un totale di 77 osservazioni per ciascun titolo.

Si calcolino i log-rendimenti di ogni singolo titolo. La media e la varianza campionaria dei log-rendimenti sono riassunte nella seguente tabella e rappresentate nel successivo grafico E-V :

	Intesa Sanpaolo	Unicredit	Fiat	Mediaset
medie	0,00164	0,00140	-0,00036	0,00445
var	0,00052	0,00049	0,00054	0,00078
sqm	0,02270	0,02208	0,02329	0,02798



La matrice Σ di varianze e covarianze dei 4 titoli azionari può essere stimata con i dati campionari, così come la matrice inversa Σ^{-1} :

$$\Sigma =$$

	Intesa San Paolo	Unicredit	Fiat	Mediaset
Intesa San Paolo	0,00052	0,00042	0,00030	0,00023
Unicredit	0,00042	0,00049	0,00032	0,00018
Fiat	0,00030	0,00032	0,00054	0,00015
Mediaset	0,00023	0,00018	0,00015	0,00078

$$\Sigma^{-1} =$$

	Intesa San Paolo	Unicredit	Fiat	Mediaset
Intesa San Paolo	6913,57	-5375,30	-507,75	-671,94
Unicredit	-5375,30	7660,80	-1551,65	116,66
Fiat	-507,75	-1551,65	3077,59	-106,82
Mediaset	-671,94	116,66	-106,82	1465,30

Si hanno ora tutti i dati a disposizione per calcolare gli scalari A, B, C, D :

$$A = \mu' \Sigma^{-1} \bar{\mathbf{1}} = \bar{\mathbf{1}}' \Sigma^{-1} \mu = 5,02$$

$$B = \mu' \Sigma^{-1} \mu = 0,03$$

$$C = \bar{\mathbf{1}}' \Sigma^{-1} \bar{\mathbf{1}} = 2923,64$$

$$D = BC - A^2 = 69,69$$

che portano alla stima del vettore di pesi ω_P cercato :

$$\omega_P = g + \mu_P h$$

dove :

$$g = \frac{B \Sigma^{-1} \bar{1} - A \Sigma^{-1} \mu}{D} = \frac{B}{D} \Sigma^{-1} \bar{1} - \frac{A}{D} \Sigma^{-1} \mu = \begin{bmatrix} 0,092 \\ 0,184 \\ 0,755 \\ -0,031 \end{bmatrix}$$

$$h = \frac{C \Sigma^{-1} \mu - A \Sigma^{-1} \bar{1}}{D} = \frac{C}{D} \Sigma^{-1} \mu - \frac{A}{D} \Sigma^{-1} \bar{1} = \begin{bmatrix} 17,66 \\ 62,27 \\ -257,79 \\ 177,86 \end{bmatrix}$$

Quindi :

$$\omega_1 = 0,092 + 17,66 \mu_P$$

$$\omega_2 = 0,184 + 62,27 \mu_P$$

$$\omega_3 = 0,755 - 257,79 \mu_P$$

$$\omega_4 = -0,031 + 177,86 \mu_P$$

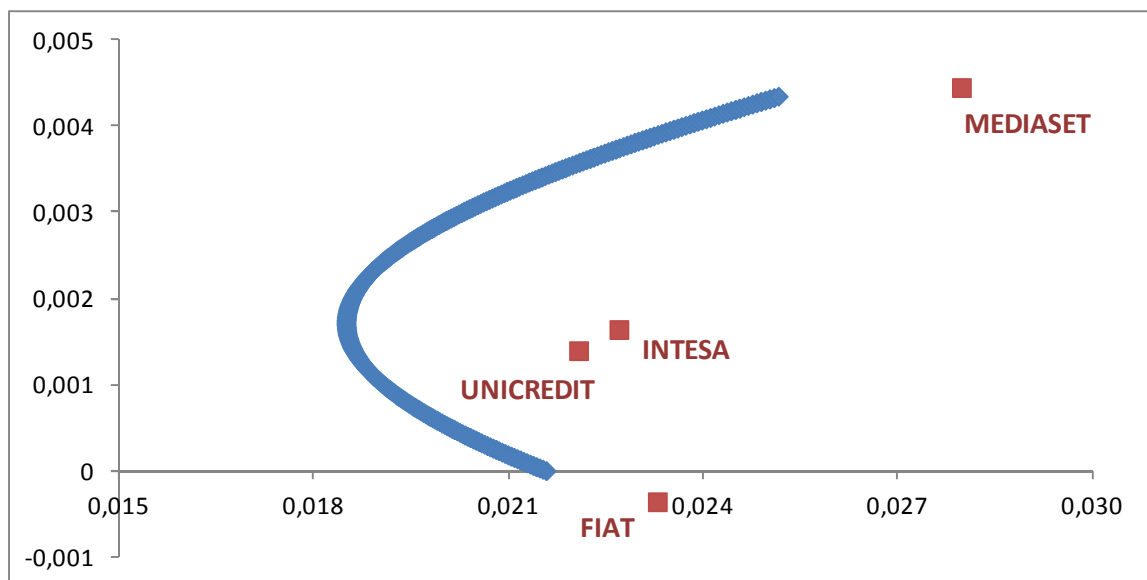
Il sistema di pesi, con somma unitaria, porta al portafoglio con varianza minima tra tutti i portafogli con rendimento atteso pari a μ_P .

Per μ_P che varia da 0 al max $[\mu_i] = 0,00455$ si ottengono i possibili portafogli efficienti riassunti nella seguente tabella⁷ :

⁷ In generale, per la proprietà di Cauchy della media, il rendimento atteso di un portafoglio varia tra il minimo e il massimo rendimento atteso dei titoli che lo compongono, ma nel caso in cui il minimo risulti negativo (come Fiat nell'esempio) si pone il minimo di $\mu_P = 0$ poiché nessun investitore allocherebbe alcuna risorsa in un portafoglio con rendimento atteso negativo.

μ_P	ω_1	ω_2	ω_3	ω_4	σ^2_P	σ_P
0,0000	0,0923	0,1841	0,7540	- 0,0305	0,0005	0,0216
0,0001	0,0934	0,1879	0,7382	- 0,0195	0,0005	0,0214
0,0002	0,0951	0,1936	0,7144	- 0,0031	0,0004	0,0211
0,0003	0,0970	0,2003	0,6867	0,0160	0,0004	0,0208
0,0004	0,0986	0,2061	0,6629	0,0324	0,0004	0,0205
0,0005	0,1005	0,2128	0,6352	0,0515	0,0004	0,0202
0,0006	0,1021	0,2185	0,6114	0,0679	0,0004	0,0200
0,0007	0,1040	0,2252	0,5837	0,0871	0,0004	0,0197
0,0008	0,1056	0,2309	0,5600	0,1035	0,0004	0,0195
0,0009	0,1075	0,2376	0,5322	0,1226	0,0004	0,0193
0,0010	0,1092	0,2434	0,5085	0,1390	0,0004	0,0191
...	...	...	...	...	...	...
0,0040	0,1621	0,4299	- 0,2639	0,6719	0,0006	0,0235
0,0041	0,1640	0,4366	- 0,2916	0,6910	0,0006	0,0239
0,0042	0,1656	0,4424	- 0,3154	0,7074	0,0006	0,0243
0,0043	0,1675	0,4491	- 0,3431	0,7265	0,0006	0,0248
0,0044	0,1691	0,4548	- 0,3669	0,7429	0,0006	0,0252
0,0045	0,1709	0,4610	- 0,3926	0,7607	0,0007	0,0256

La *frontiera efficiente* è data dall'insieme di tutti i portafogli efficienti, e può essere visualizzata come una curva in un grafico E-V :

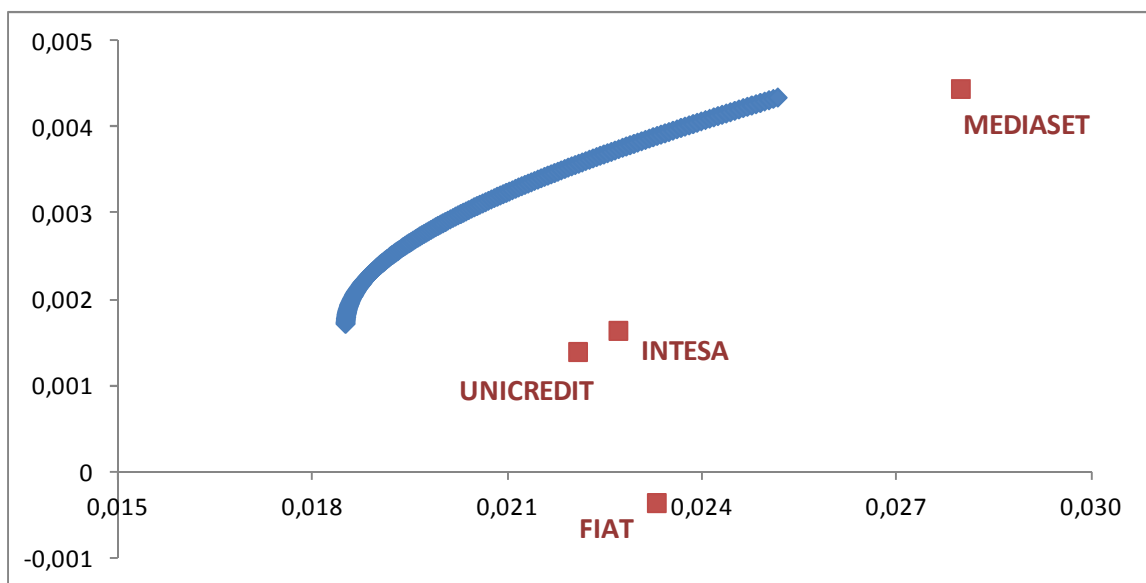


Tutti i punti all'interno di tale curva, dove risiedono i titoli stessi, sono portafogli possibili ma non efficienti, poiché a parità di σ_P avranno sicuramente un portafoglio con rendimento atteso superiore, oppure a parità di μ_P avranno sicuramente un portafoglio con rischio minore, quindi in definitiva avranno sempre uno o più portafogli (quelli sulla frontiera) che li dominano nel grafico.

Tutti i punti esterni a tale curva, invece, non sono portafogli possibili, poiché la frontiera indica proprio i portafogli che minimizzano il rischio a parità di μ_P , che equivale a massimizzare il rendimento atteso a parità di σ_P .

Tra tutti i punti che giacciono sulla frontiera, però, anche quelli appartenenti alla parte bassa della curva, sotto al punto di minimo scarto quadratico medio, non sono efficienti, poiché a parità di rischio hanno un altro portafoglio con rendimento superiore (sopra di essi nella curva).

In definitiva, il portafoglio con allocazione ottimale va scelto dunque tra tutti i portafogli possibili ed efficienti appartenenti alla seguente semi-curve :



E' compito dell'investitore decidere il portafoglio ottimale tra questi, dato un prefissato livello di rendimento atteso μ_P , ma tipicamente sono due i portafogli su cui un decisore pone la sua attenzione :

- il portafoglio a varianza minima (Global Minimum Variance Portfolio, GMVP)
- il portafoglio tangente (Tangency Portfolio)

1.7 Portafoglio a varianza minima

Il Global Minimum Variance Portfolio è visualizzabile graficamente come il “punto di svolta” della curva di frontiera dei portafogli efficienti.

E' quindi il portafoglio con la varianza minima tra tutti i portafogli efficienti.

Esso è quel portafoglio che minimizza :

$$\sigma_P^2 = \omega_P' \Sigma \omega_P$$

dove :

$$\omega_P = g + h \mu_P$$

Quindi :

$$\begin{aligned}\sigma_P^2 &= (g + h \mu_P)' \Sigma (g + h \mu_P) = \\ &= g' \Sigma g + 2 g' \Sigma h \mu_P + h' \Sigma h \mu_P^2\end{aligned}$$

che si minimizza ponendo :

$$\frac{\partial \sigma_P^2}{\partial \mu_P} = 2 g' \Sigma h + 2 h' \Sigma h \mu_P = 0$$

Il cui rendimento atteso è pari a :

$$\mu_{P (min.)} = - \frac{g' \Sigma h}{h' \Sigma h}$$

con varianza pari a :

$$\sigma_{P (min.)}^2 = g' \Sigma g - \frac{(g' \Sigma h)^2}{h' \Sigma h}$$

Nell'esempio 2 :

$$\mu_{P (min.)} = 0,00171795 \quad \sigma_{P (min.)}^2 = 0,00034204 \quad \sigma_{P (min.)} = 0,0184943$$

e il sistema di pesi associato è :

$$\omega_{P (min.)} = g + h \mu_{P (min.)} = [0,123 \quad 0,291 \quad 0,312 \quad 0,275]'$$

Si noti come il sistema di pesi associato al portafoglio a varianza minima, può essere stimato anche indipendentemente dal livello di μ_P , ma prendendo in considerazione solo ed esclusivamente la matrice Σ di varianze e covarianze.

Ciò è possibile poiché il sistema di pesi che annulla la derivata prima della funzione lagrangiana è dato da :

$$\omega_P = \lambda_1 \Sigma^{-1} \mu + \lambda_2 \Sigma^{-1} \mathbf{1}$$

Ma, come abbiamo già visto :

$$\lambda_1 = \frac{-A + C \mu_P}{D}$$

$$\lambda_2 = \frac{B - A \mu_P}{D}$$

Quindi :

$$\begin{aligned} \omega_P &= \left(\frac{-A + C \mu_P}{D} \right) \Sigma^{-1} \mu + \left(\frac{B - A \mu_P}{D} \right) \Sigma^{-1} \mathbf{1} = \\ &= \frac{(C \mu_P - A) \Sigma^{-1} \mu + (B - A \mu_P) \Sigma^{-1} \mathbf{1}}{D} \end{aligned}$$

Ma si dimostra⁸ che il portafoglio a varianza minima si ottiene quando : $\mu_P = \frac{A}{C}$

Quindi :

$$\omega_{P (min.)} = \frac{\left(C \left(\frac{A}{C} \right) - A \right) \Sigma^{-1} \mu + \left(B - A \left(\frac{A}{C} \right) \right) \Sigma^{-1} \mathbf{1}}{D} =$$

⁸ Per la dimostrazione si rimanda a Merton (1972), in cui :

$$\sigma_P^2 = \lambda_1 \mu_P + \lambda_2$$

Quindi :

$$\sigma_P^2 = \frac{-A+C \mu_P}{D} \mu_P + \frac{B-A \mu_P}{D} = \frac{-A+C \mu_P^2+B-A \mu_P}{D} = \frac{C \mu_P^2-2 A \mu_P+B}{D}$$

che è un polinomio di ordine 2, che si minimizza ponendo :

$$\frac{\partial \sigma_P^2}{\partial \mu_P} = \frac{2 C \mu_P - 2 A}{D} = 0$$

Ovvero :

$$C \mu_P - A = 0$$

che porta a : $\mu_P = \frac{A}{C}$

$$= \frac{0 + (B - \frac{A^2}{C})\Sigma^{-1}\bar{1}}{BC - A^2} = \frac{(\frac{BC - A^2}{C})\Sigma^{-1}\bar{1}}{BC - A^2} = \frac{\Sigma^{-1}\bar{1}}{C} = \frac{\Sigma^{-1}\bar{1}}{\bar{1}'\Sigma^{-1}\bar{1}}$$

In tal caso, si dimostra⁹ quindi che, se $\omega_P (min.) = \frac{\Sigma^{-1}\bar{1}}{C}$, la varianza del portafoglio a varianza minima, è data da :

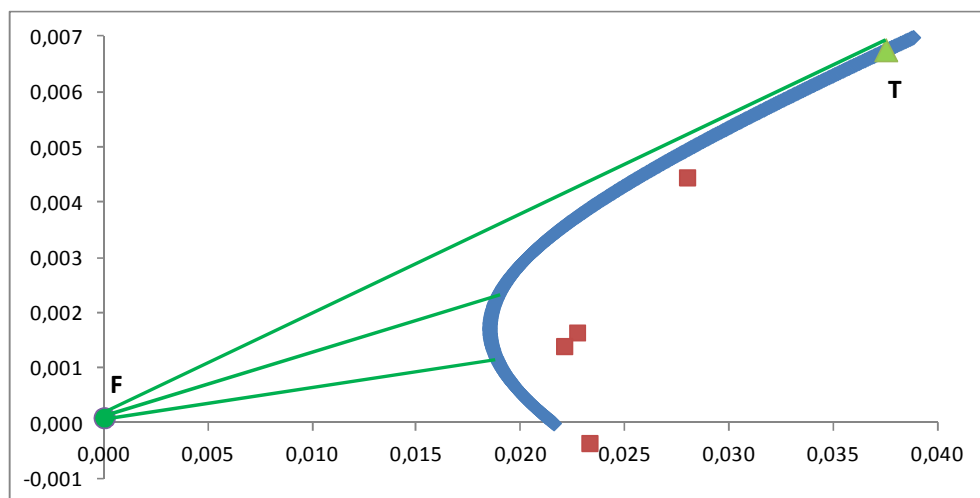
$$\sigma_P^2 = \omega_P' \Sigma \omega_P = \frac{1}{\bar{1}'\Sigma^{-1}\bar{1}} = \frac{1}{C}$$

Tale risultato è di notevole importanza poiché mostra come il rischio del portafoglio a varianza minima dipenda esclusivamente dagli elementi che compongono la matrice di varianze e covarianze, la cui stima risulta dunque cruciale nell'analisi in oggetto.

1.8 Portafoglio tangente

Si introduca ora la possibilità di inserire nel portafoglio un titolo *risk-free*, ovvero dal rendimento privo di rischio,¹⁰ con rendimento atteso pari a μ_F .

Il punto F in figura, rappresenta il valore di (σ_F, μ_F) del titolo *risk-free* ponendo $\mu_F = 0,0001$ (il valore in ascissa è dunque pari a $\sigma_F = 0$)



⁹ si veda Merton (1972)

¹⁰ si veda Ruppert (2011)

I possibili valori di (σ_P, μ_P) che contengano sia il portafoglio (efficiente) sia il nuovo titolo *risk-free* sono rappresentati dai segmenti che congiungono il punto F alla curva di frontiera. Tra tutti i segmenti possibili, quello che congiunge F al punto T risiede graficamente più in alto rispetto a tutti gli altri, quindi a parità di rischio σ gli corrisponde un rendimento atteso più elevato.

In finanza, la retta che nel grafico rischio-rendimento ha come intercetta il rendimento atteso μ_F del titolo risk-free e come pendenza il cosiddetto indice di Sharpe¹¹, è nota come Security Market Line (SML).

Il portafoglio T è dunque quel portafoglio efficiente che massimizza il valore dell'indice di Sharpe, ovvero che massimizza la pendenza della Security Market Line.

Tale portafoglio viene chiamato “portafoglio tangente”, e ha rendimento atteso pari a :

$$\mu_P = \boldsymbol{\omega}' \boldsymbol{\mu} + (1 - \boldsymbol{\omega}' \bar{\mathbf{1}}) \mu_F$$

Ciò che viene minimizzata è sempre la varianza del portafoglio :

$$\min \quad \boldsymbol{\omega}' \boldsymbol{\Sigma} \boldsymbol{\omega}$$

$$\text{sub} \quad \boldsymbol{\omega}' \boldsymbol{\mu} + (1 - \boldsymbol{\omega}' \bar{\mathbf{1}}) \mu_F = \mu_P$$

Da cui la funzione lagrangiana diventa :

$$L(\boldsymbol{\omega}, \delta_1) = \boldsymbol{\omega}' \boldsymbol{\Sigma} \boldsymbol{\omega} + \delta [\mu_P - \boldsymbol{\omega}' \boldsymbol{\mu} - (1 - \boldsymbol{\omega}' \bar{\mathbf{1}}) \mu_F]$$

Se si pone la derivata prima uguale a 0 :

$$\frac{\partial L(\boldsymbol{\omega}, \delta_1)}{\partial \boldsymbol{\omega}} = 2 \boldsymbol{\Sigma} \boldsymbol{\omega} + \delta (-\boldsymbol{\mu} + \bar{\mathbf{1}} \mu_F) = 0$$

si ottiene :

$$\boldsymbol{\omega}_P = \lambda \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \mu_F \bar{\mathbf{1}})$$

¹¹ letteralmente “Sharpe ratio”, da William Sharpe (1999) che è dato da $\frac{\mu_P - \mu_F}{\sigma_P}$

dove $\lambda = \delta/2$ è il nuovo moltiplicatore di Lagrange, che si determina sostituendo ω_P nel vincolo :

$$\omega'_P \mu + (1 - \omega'_P \bar{\mathbf{1}}) \mu_F = \mu_P$$

$$\omega'_P (\mu - \mu_F \bar{\mathbf{1}}) = \mu_P - \mu_F$$

$$\lambda (\mu - \mu_F \bar{\mathbf{1}})' \Sigma^{-1} (\mu - \mu_F \bar{\mathbf{1}}) = \mu_P - \mu_F$$

che porta a :

$$\lambda = \frac{\mu_P - \mu_F}{(\mu - \mu_F \bar{\mathbf{1}})' \Sigma^{-1} (\mu - \mu_F \bar{\mathbf{1}})}$$

Quindi :

$$\omega_P = \lambda \bar{\omega}$$

dove :

$$\bar{\omega} = \Sigma^{-1} (\mu - \mu_F \bar{\mathbf{1}})$$

Il vettore di pesi $\bar{\omega}$ così ottenuto non rappresenta un'allocazione di portafoglio, poiché la somma dei pesi che lo compongono non è pari a 1.

Se si pone la condizione che $\bar{\omega}' \bar{\mathbf{1}} = 1$ si ottiene il portafoglio tangente :

$$\omega_T = \frac{\bar{\omega}}{\bar{\mathbf{1}}' \bar{\omega}}$$

Il valore del rendimento atteso e della varianza del portafoglio tangente, possono essere stimati sostituendo i pesi ω_T all'interno delle formule :

$$\mu_T = \omega'_T \mu$$

$$\sigma_T^2 = \omega'_T \Sigma \omega_T$$

Dall'analisi grafica si nota che il portafoglio tangente può essere stimato mettendo a sistema l'equazione della retta tangente con l'equazione della curva di frontiera, in modo tale da stimare il rendimento atteso e la varianza di T indipendentemente dal valore dei parametri ω_T .

Per fare ciò¹², senza perdita di generalità, si ponga¹³ $\mu_F = 0$.

In tal caso, l'equazione della retta di tangenza, posto :

$$SR_P = \text{Sharpe ratio} = \frac{\mu_P}{\sigma_P}$$

diventa : $\sigma_P = \frac{\mu_P}{SR_P}$

da cui : $\sigma_P^2 = \frac{\mu_P^2}{SR_P^2}$

L'equazione della curva di frontiera, può essere trovata a partire dalle formule mostrate in precedenza (paragrafo 1.6), in cui :

$$\omega_P = \lambda_1 \Sigma^{-1} \mu + \lambda_2 \Sigma^{-1} \bar{1} = \Sigma^{-1} (\lambda_1 \mu + \lambda_2 \bar{1})$$

dove :

$$\lambda_1 = \frac{C \mu_P - A}{D}$$

$$\lambda_2 = \frac{B - A \mu_P}{D}$$

Quindi :

$$\begin{aligned} \sigma_P^2 &= \omega'_P \Sigma \omega_P = \\ &= \omega'_P \Sigma \Sigma^{-1} (\lambda_1 \mu + \lambda_2 \bar{1}) = \\ &= \lambda_1 \omega'_P \mu + \lambda_2 \omega'_P \bar{1} = \\ &= \lambda_1 \mu_P + \lambda_2 \end{aligned}$$

¹² si veda "Modello di simulazione per un portafoglio diversificato" Palomba (2004).

¹³ i risultati mostrati saranno facilmente estendibili al generico caso in cui $\mu_F \neq 0$, sostituendo a μ_P il termine $(\mu_P - \mu_F)$

Mettendo a sistema le equazioni :

$$\begin{cases} \sigma_P^2 = \lambda_1 \mu_P + \lambda_2 \\ \sigma_P^2 = \frac{\mu_P^2}{SR_P^2} \end{cases}$$

Si ottiene :

$$\lambda_1 \mu_P + \lambda_2 = \frac{\mu_P^2}{SR_P^2}$$

$$\frac{C \mu_P - A}{D} \mu_P + \frac{B - A \mu_P}{D} = \frac{\mu_P^2}{SR_P^2}$$

$$\frac{C \mu_P^2 - 2A\mu_P + B}{D} = \frac{\mu_P^2}{SR_P^2}$$

$$(C \mu_P^2 - 2A \mu_P + B) SR_P^2 = D \mu_P^2$$

$$C \mu_P^2 SR_P^2 - 2A \mu_P SR_P^2 + B \cdot SR_P^2 - D \mu_P^2 = 0$$

$$(C \cdot SR_P^2 - D) \mu_P^2 - 2A \cdot SR_P^2 \mu_P + B \cdot SR_P^2 = 0$$

Che, essendo un'equazione di secondo grado in μ_P , si risolve ponendo :

$$4 A^2 \cdot SR_P^4 - 4 B \cdot SR_P^2 (C \cdot SR_P^2 - D) = 0$$

Ma, essendo $SR_P^2 \neq 0$

$$A^2 \cdot SR_P^2 - B (C \cdot SR_P^2 - D) = 0$$

$$A^2 \cdot SR_P^2 - BC \cdot SR_P^2 - BD = 0$$

$$(A^2 - BC) \cdot SR_P^2 - BD = 0$$

$$D \cdot SR_P^2 - BD = 0$$

$$SR_P^2 = B$$

$$SR_P = +\sqrt{B}$$

Da cui, sostituendo nelle formule precedenti, si ottiene :

$$\mu_T = \frac{B}{A} \qquad \sigma_T^2 = \frac{B}{A^2}$$

Le stime di μ_T e σ_T per il generico caso in cui $\mu_F \neq 0$, si ottengono sostituendo $(\boldsymbol{\mu} - \mu_F)$ a $\boldsymbol{\mu}$ nelle stime di A e B, ottenendo :

$$\mu_T = \frac{B}{A} = \frac{\boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \mu_F)}{\bar{\mathbf{1}}' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \mu_F)}$$

$$\sigma_T^2 = \frac{B}{A^2} = \frac{(\boldsymbol{\mu} - \mu_F)' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \mu_F)}{(\bar{\mathbf{1}}' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \mu_F))^2}$$

Nell'esempio precedente, per ottenere un portafoglio T tangente alla semi-curva dei portafogli efficienti, si può porre un titolo *risk-free* con rendimento atteso $\mu_F = 0,0001$.

In tal caso :

$$\bar{\boldsymbol{\omega}} = \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \mu_F \bar{\mathbf{1}}) = [1,0011 \quad 2,8603 \quad -4,6703 \quad 5,5392]'$$

$$\boldsymbol{\omega}_T = \frac{\bar{\boldsymbol{\omega}}}{\bar{\mathbf{1}}' \bar{\boldsymbol{\omega}}} = [0,211 \quad 0,605 \quad -0,987 \quad 1,171]'$$

che portano ad un portafoglio con media e varianza pari a :

$$\mu_{P(Tan.)} = 0,0067 \qquad \sigma_{P(Tan.)}^2 = 0,0014 \qquad \sigma_{P(Tan.)} = 0,0375$$

come visualizzato in precedenza nel grafico E-V.

Dal valore dei pesi $\boldsymbol{\omega}_T$ ottenuto in questo esempio e soprattutto dal valore dello scarto quadratico medio $\sigma_{P(Tan.)}$, che risulta essere maggiore dei singoli scarti quadratici medi dei titoli che compongono il portafoglio, si nota come non sempre il portafoglio tangente risulti un portafoglio ottimale per un investitore, in quanto non sempre la curva di frontiera ottenuta tramite il “vecchio” portafoglio (in assenza del titolo *risk-free*) ammette una retta tangente in corrispondenza del punto $(0, \mu_F)$ con $\mu_F \geq 0$.

1.9 Assenza di short sellings

Quanto visto finora, prevedeva la possibilità di effettuare vendite allo scoperto da parte dell'investitore poiché non si escludeva la possibilità che alcuni pesi ω_i potessero essere anche negativi, ovvero nella minimizzazione del rischio l'unico vincolo posto ai parametri ω_i era che la loro somma fosse unitaria.

Imporre un nuovo vincolo tale per cui ogni ω_i sia ≥ 0 è un'operazione che richiede di soddisfare le cosiddette condizioni di Kuhn-Tucker nella minimizzazione vincolata¹⁴, ma è talvolta necessario nel caso (verosimile nella realtà) in cui l'investitore non abbia la possibilità di effettuare vendite allo scoperto.

In tal caso, per trovare i portafogli di frontiera efficiente, si tratta di minimizzare :

$$\min \quad \boldsymbol{\omega}' \boldsymbol{\Sigma} \boldsymbol{\omega}$$

sotto i vincoli :

$$\text{sub} \quad \boldsymbol{\omega}' \boldsymbol{\mu} = \mu_P$$

$$\boldsymbol{\omega}' \mathbf{1} = 1$$

$$\boldsymbol{\omega} \geq \bar{\mathbf{0}}$$

dove $\bar{\mathbf{0}}$ è un vettore-colonna composto da N zeri.

Tale minimizzazione non può essere risolta tramite il classico metodo Lagrangiano, a causa proprio dei vincoli di disuguaglianza, ed è perciò necessario ricorrere a metodi di programmazione quadratica, che sono implementati in molti pacchetti di software avanzati come Matlab ed R.

Tutti i problemi di programmazione quadratica hanno come scopo quello di minimizzare una funzione-obiettivo quadratica, che nel nostro caso sarà $\boldsymbol{\omega}' \boldsymbol{\Sigma} \boldsymbol{\omega}$, sotto vincoli lineari di uguaglianza e disuguaglianza.

I vincoli lineari di disuguaglianza sono dati da :

$$\mathbf{A} \boldsymbol{\omega} \leq \mathbf{b}$$

¹⁴ Per le quali si rimanda a Elton, Gruber (1981)

dove $\mathbf{A}$ è una matrice ($m \times N$) e $\mathbf{b}$ è un vettore ($m \times 1$), in cui m è il numero di vincoli di disuguaglianza.

Nel caso in oggetto, i vincoli di disuguaglianza sono N , uno per ogni ω_i , quindi la matrice $\mathbf{A}$ corrisponde a una matrice identità ($N \times N$) negativa :

$$\mathbf{A} = -\mathbf{I} = \begin{bmatrix} -1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 \end{bmatrix}$$

e il vettore $\mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ è un vettore-colonna di N zeri. Cosicché : $-\omega \leq 0$

Similmente, i vincoli lineari di uguaglianza sono dati genericamente da :

$$\mathbf{A}_{eq} \omega = \mathbf{b}_{eq}$$

dove $\mathbf{A}_{eq}$ è una matrice ($p \times N$) e $\mathbf{b}_{eq}$ è un vettore ($p \times 1$), in cui p è il numero di vincoli di uguaglianza.

I vincoli di uguaglianza sono due, quindi :

$$\mathbf{A}_{eq} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \mu_1 & \mu_2 & \dots & \mu_N \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{1}}' \\ \boldsymbol{\mu}' \end{bmatrix}$$

$$\mathbf{b}_{eq} = \begin{bmatrix} 1 \\ \mu_p \end{bmatrix}$$

Da cui :

$$\begin{bmatrix} \bar{\mathbf{1}}' \\ \boldsymbol{\mu}' \end{bmatrix} \omega = \begin{bmatrix} 1 \\ \mu_p \end{bmatrix}$$

Nel paragrafo 2.3 verrà mostrata, tramite un'applicazione, la procedura di programmazione quadratica che permette la stima del portafoglio tangente e del portafoglio a varianza minima in assenza di short sellings.

1.10 Stime alternative della matrice di varianze e covarianze

Nel corso della trattazione, è stata più volte messa in luce l'importanza della stima della matrice Σ di varianze e covarianze tra i rendimenti dei titoli azionari per individuare i portafogli efficienti.

I metodi di stima della matrice di varianze e covarianze sono molteplici e, nonostante l'utilizzo delle varianze e covarianze campionarie sembrerebbe la scelta più semplice e scontata (nonché quella utilizzata negli esempi mostrati fino ad ora), un'ampia letteratura riconosce che potrebbe non essere il metodo migliore.

I limiti di impiego della matrice di varianze e covarianze campionaria possono essere facilmente evidenziati se si prende in considerazione il portafoglio a varianza minima, che, come abbiamo visto in precedenza, è l'unico portafoglio sulla frontiera efficiente calcolabile senza l'utilizzo dei rendimenti attesi dei singoli titoli.

I pesi associati al portafoglio a varianza minima dipendono infatti esclusivamente dalla matrice di varianze e covarianze Σ e sono dati da :

$$\omega_{P(min.)} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}$$

Prendendo in considerazione l'Esempio 2, la cui matrice di varianze e covarianze campionaria è definita a pagina 26, il portafoglio a varianza minima è stato trovato come :

$$\omega_{P(min.)} = [0,123 \quad 0,291 \quad 0,312 \quad 0,275]'$$

che porta a un rendimento atteso e a una varianza pari a :

$$\mu_{P(min.)} = 0,00171795$$

$$\sigma_{P(min.)}^2 = 0,00034204$$

Essendo il vettore μ delle medie dei singoli rendimenti dato da :

	Intesa Sanpaolo	Unicredit	Fiat	Mediaset
medie	0,00164	0,00140	-0,00036	0,00445

significa che il portafoglio a varianza minima, ritenuto ovviamente ottimale perché efficiente e a minimo rischio tra tutti quelli efficienti, pone il peso maggiore $\omega_3 = 31,2\%$ al titolo “Fiat” che è l’unico titolo in portafoglio ad avere un rendimento atteso negativo!

Le ragioni di ciò stanno evidentemente nella conformazione della matrice Σ , che è l’unico elemento entrato in gioco per il calcolo dei pesi $\omega_{P(min.)}$.

Inoltre, dalla matrice di varianze e covarianze campionaria Σ , deriva la seguente matrice di correlazione tra i rendimenti dei titoli :

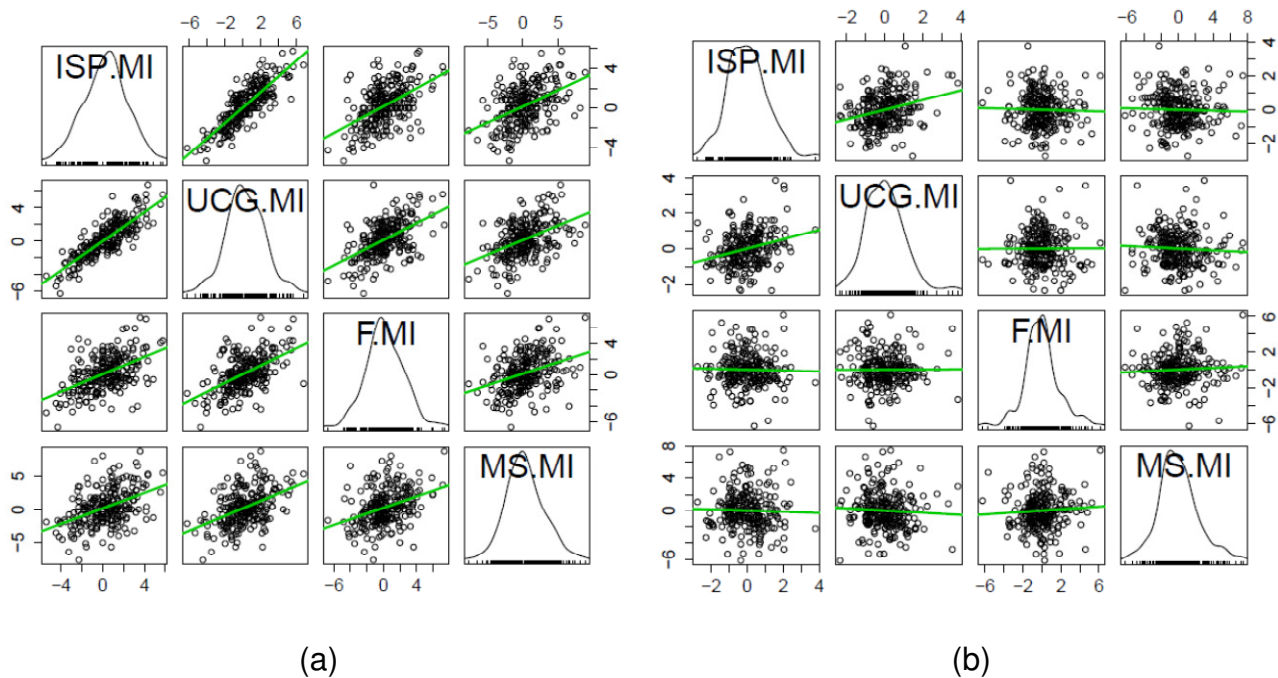
	Intesa	Unicredit	Fiat	Mediaset
Intesa Sanpaolo	1			
Unicredit	0,841	1		
Fiat	0,584	0,625	1	
Mediaset	0,356	0,286	0,235	1

da cui si può osservare come tutti i titoli, appartenenti a diversi settori merceologici, abbiano correlazione positiva e i titoli appartenenti a gruppi bancari (Intesa ed Unicredit) siano effettivamente altamente correlati, come avremmo potuto aspettarci nella realtà.

Tale risultato può essere riconducibile ad un fenomeno di correlazione “spuria”, ovvero una correlazione che si manifesta tra due variabili ma dovuta ad una terza (o più) variabile estranea ad esse, le quali nella realtà non dovrebbero risultare correlate.

Nel mercato finanziario, la presenza di un’alta correlazione positiva all’interno della matrice di varianze e covarianze campionaria è dovuta, come nel caso di Fiat e Unicredit nell’esempio precedente, al fatto che i titoli appartengano tutti allo stesso mercato finanziario, in questo caso rappresentato dall’indice FTSEMIB, il cui incremento/decremento porta ad un movimento simultaneo di tutti i titoli che lo compongono.

Per mostrare tale limite si consideri il seguente grafico che rappresenta le correlazioni parziali fra i quattro titoli : Intesa San Paolo (ISP.MI), Unicredit (UCG.MI), Fiat (F.MI) e Mediaset (MS.MI).



Nella prima figura (a) vengono visualizzate le correlazioni (campionarie) osservate tra i 4 titoli nell'arco temporale di un anno. Come si nota, tutte le correlazioni risultano positive.

Depurando però l'effetto di mercato in figura (b), ossia prendendo in considerazione il vettore dei residui ottenuto regredendo ciascun titolo in funzione dell'indice FTSEMIB, si nota come le correlazioni parziali tra i quattro titoli nell'arco dello stesso anno tendono a smorzarsi, fino a una correlazione pressoché nulla tra un titolo che non appartiene al settore bancario (Fiat o Mediaset) e i due titoli Intesa San Paolo e Unicredit, come ci si sarebbe potuti aspettare nella realtà, mentre rimane positiva la correlazione tra i due titoli bancari.

L'esempio dunque mette in luce come la stima delle correlazioni tra i titoli azionari appartenenti ad uno stesso mercato, possono essere influenzate anche dalla semplice appartenenza dei titoli al mercato stesso. Quindi l'utilizzo delle stime provenienti dal campione portano in alcuni casi all'ottenimento di portafogli ottimali non plausibili, che hanno sviluppato, nel corso degli anni, l'esigenza di una varietà di tecniche alternative per la stima della matrice di varianze e covarianze tra i titoli. Tra le più utilizzate, oltre alla stima di massima verosimiglianza della matrice inversa di varianze e covarianze, cui si accennerà nel seguito, si annoverano anche il "modello a un fattore"¹⁵ ed il "modello a correlazione costante"¹⁶.

¹⁵ Benninga (2011)

¹⁶ Elton e Grumber (1973)

Il *modello a un fattore* (*Single Index Model, SIM*) si basa sull'assunto che il rendimento di ogni singolo titolo $i=1,2,...,N$ può essere espresso tramite un modello di regressione lineare in funzione di un determinato indice di mercato X :

$$R_i = \alpha_i + \beta_i R_x + \varepsilon_i$$

dove la correlazione tra i residui ε_i ed ε_j , relativi rispettivamente al titolo i e al titolo j , è nulla. Sulla base di questa ipotesi, si ha che :

$$\sigma_{ij} = \begin{cases} \beta_i \beta_j \sigma_x^2 & se \quad i \neq j \\ \sigma_i^2 & se \quad i = j \end{cases}$$

quindi la stima della varianza di ogni rendimento rimane immutata, lasciando inalterato il rischio intrinseco di ogni titolo.

Il *modello a correlazione costante*, invece, stima la matrice di varianze e covarianze sotto l'assunto che tutte le covarianze facciano riferimento allo stesso coefficiente di correlazione ρ , che generalmente è posto pari al coefficiente di correlazione medio dei titoli. Dal momento che : $Cov(R_i, R_j) = \sigma_{ij} = \rho_{ij} \sigma_i \sigma_j$, si ottiene :

$$\sigma_{ij} = \begin{cases} \rho \sigma_i \sigma_j & se \quad i \neq j \\ \sigma_i^2 & se \quad i = j \end{cases}$$

Anche in questo caso non si applicano modifiche alle stime delle varianze campionarie, lasciando inalterata la diagonale principale della matrice.

Una classe di metodi che ha recentemente conquistato popolarità per la stima della matrice di varianze e covarianze sono i cosiddetti metodi *shrinkage*, in cui la matrice **S** di varianze e covarianze campionaria viene *penalizzata* (da qui il termine shrinkage).

1.11 Metodi di stima *shrinkage*

Numerosi autori, in particolare Daniels e Kass (2001), proposero una stima penalizzata degli autovalori di una matrice di varianze e covarianze, al fine di assicurarne una stima maggiormente stabile, poiché spesso gli autovalori campionari risultano essere distorti, con valori troppo piccoli o troppo elevati.

Ma è con i lavori di Ledoit e Wolf (2003) che per la prima volta i metodi di penalizzazione vengono applicati direttamente ad una matrice di varianze e covarianze di rendimenti in ambito finanziario. Ai fini di un'ottimizzazione di portafoglio efficiente, gli autori suggeriscono di utilizzare una matrice ottenuta dalla matrice di varianze e covarianze campionaria dopo aver effettuato una penalizzazione. Ciò permette di far tendere gli elementi più elevati della matrice (in modulo) verso valori minori, in modo tale da ridurre sistematicamente l'errore di stima dove è maggiormente presente.

L'obiettivo è quindi quello di ottenere una stima migliore della matrice Σ , che rappresenti il giusto trade-off tra la forte distorsione di una matrice proveniente da un modello "SIM" e gli errori di stima di una matrice proveniente dal campione, e che quindi sia una giusta via di mezzo tra i due metodi.

In questo caso, la matrice di varianze e covarianze viene stimata, come nella maggior parte dei metodi di shrinkage presenti nella letteratura moderna, effettuando una combinazione lineare che dia un peso pari a $(1-\lambda)$ alla matrice $\mathbf{S}$ campionaria (dove λ è un parametro che controlla l'intensità della penalizzazione compreso tra 0 e 1) e peso λ ad una seconda (o più) matrice $\mathbf{\Gamma}$, chiamata *shrinkage target matrix*, simmetrica, definita positiva e meno soggetta a errori di stima :

$$\Sigma_{\lambda} = (1 - \lambda) \mathbf{S} + \lambda \mathbf{\Gamma}$$

Numerosi sono gli autori che nel corso degli anni più recenti hanno proposto una diversa *shrinkage target matrix* da applicare alla metodologia, tra cui si menzionano Kourtis, Dotsis e Markellos (2012) i quali propongono la matrice identità $\mathbf{I}_{(N \times N)}$, dove N è il numero di titoli azionari, in alternativa alle matrici provenienti da un modello a un fattore o a correlazione costante.

Kourtis, Dotsis e Markellos (2012) sono stati anche i primi autori a introdurre questi metodi di stima penalizzata non solo alla matrice di varianze e covarianze, ma anche alla sua inversa Θ , la cui importanza verrà chiarita nei prossimi capitoli del presente elaborato.

Molti autori propongono l'utilizzo dei metodi di stima penalizzati soprattutto nel caso in cui il numero di osservazioni campionarie T è molto inferiore al numero di assets N , e la matrice Σ di varianze e covarianze risulta essere quasi singolare. Come nel caso di De Miguel et al. (2013), in cui si focalizza l'attenzione soprattutto sulla stima ottimale del parametro di penalizzazione λ .

Una stima del parametro λ viene proposta anche da Theiler (2012) che pone tale stima inversamente proporzionale al numero di osservazioni campionarie T , poiché più T è elevato, più la stima delle covarianze campionarie approssimano il vero valore delle covarianze tra i titoli e quindi dev'essere effettuata una bassa penalizzazione alla matrice campionaria S , mentre per valori piccoli di T è necessario un alto valore di λ .

Si noti, infine, che i metodi di penalizzazione applicati alla matrice di varianze e covarianze campionaria non sono stati sviluppati nel corso degli ultimi anni solamente in ambito finanziario, ma come dimostrano i lavori di Schafer, Strimmer (2005) e Cui et al. (2005) sono molteplici i campi in cui è necessario applicare una penalizzazione alle stime ottenute tramite il campione.

Tutti i metodi di shrinkage hanno come obiettivo quello di una migliore stima della matrice Σ , effettuando una *penalizzazione* dei suoi elementi. L'idea di questo elaborato è quella di implementare tali tecniche con una metodologia che permetta di effettuare anche un'operazione di *variable selection* tra i titoli appartenenti al portafoglio, in modo tale da comprenderne e modificarne la composizione. Per ottenere ciò è possibile usufruire di una tecnica nota in letteratura come LASSO (Least Absolute Shrinkage and Selection Operator) da applicare alla matrice di varianze e covarianze di un portafoglio azionario, come verrà sviluppata nei successivi capitoli.

CAP.2

Least Absolute Shrinkage and Selection Operator

2.1 Introduzione alla tecnica LASSO

Evidenziati i limiti di impiego della matrice di varianze e covarianze campionaria, messi in luce nel capitolo precedente, obiettivo principale del presente elaborato è quello di proporre una metodologia alternativa per la stima della matrice stessa, al fine di monitorare l'andamento del portafoglio ottimo e fornire al decisore uno strumento più completo per analizzare il proprio investimento.

L'algoritmo di Least Absolute Shrinkage and Selection Operator (LASSO) trae le sue origini dagli studi di Robert Tibshirani del 1996, in cui venne proposto l'utilizzo della metrica L_1 come penalizzazione nell'ambito della regressione lineare.

Si consideri un contesto in cui :

$$y_j = \alpha + \sum_{i=1}^p \beta_i \cdot x_{ij}$$

dove la variabile Y è dipendente da p regressori X_i .

Senza perdere di generalità, si ipotizzi $\alpha = 0$ (altrimenti si interpretino y_j come scarti dalla media), quindi :

$$y_j = \sum_{i=1}^p \beta_i \cdot x_{ij}$$

La stima dei coefficienti β_i basata sul metodo dei minimi quadrati, si ottiene minimizzando la somma del quadrato degli scarti tra l'osservazione ed il modello :

$$\min_{\beta} \sum_{j=1}^n \left(y_j - \sum_{i=1}^p \beta_i \cdot x_{ij} \right)^2$$

Ci sono però molti casi in cui le stime ottenute con il metodo dei minimi quadrati non sono ritenute soddisfacenti, in particolare per due ragioni :

1. Precisione : gli stimatori OLS (Ordinary Least Squares), pur essendo non distorti, risultano essere i più efficienti nella classe degli stimatori non distorti (*Best Linear Unbiased Estimator, BLUE*) solamente nel caso in cui i residui del modello soddisfino le ipotesi di omoschedasticità ed incorrelazione, non sempre verificate.
2. Interpretazione : nel caso in cui il numero p di regressori fosse elevato, si vorrebbe spesso determinare un sottoinsieme di essi, in modo tale da individuare i predittori che più spiegano gli effetti sulla variabile Y .

Sono state sviluppate quindi diverse tecniche per implementare e migliorare le stime OLS, tra cui, appunto, la LASSO (Tibshirani, 1996), la Ridge Regression (Miller, 2002; Draper e Smith, 1998) e la Continuum Regression (Sundberg, 2002).

La Continuum Regression viene utilizzata nel caso di multicollinearità tra le variabili esplicative. Essa rappresenta un metodo di regressione strettamente correlato con molte altre tecniche di stima penalizzata, come la Ridge regression ed i metodi OLS, PCR (Principal Components Regression) e PLSR (Partial Least Squares Regression) che possono essere visti come casi particolari di Continuum Regression.

La Ridge Regression mantiene significativi tutti i predittori X_i ma con coefficienti tipicamente minori (in modulo) rispetto a quelli ottenuti con le stime OLS, introducendo nella minimizzazione una penalità in L_2 del tipo :

$$\min_{\beta} \sum_{j=1}^n \left(y_j - \sum_{i=1}^p \beta_i \cdot x_{ij} \right)^2 + \lambda \sum_{i=1}^p \beta_i^2$$

che equivale a minimizzare :

$$\begin{aligned} \min_{\beta} \sum_{j=1}^n \left(y_j - \sum_{i=1}^p \beta_i \cdot x_{ij} \right)^2 \\ \text{sub } \sum_{i=1}^p \beta_i^2 \leq t \end{aligned}$$

dove t è inversamente proporzionale a λ e $\lambda \geq 0$ rappresenta il *parametro di penalizzazione* che, se posto = 0, fa coincidere le stime di Ridge Regression con quelle OLS, mentre al crescere di λ (quindi al tendere di $t \rightarrow 0$) la stima dei parametri diminuisce, ma senza mai arrivare a zero. Quindi nessuna variabile X_i viene esclusa e il problema di *variable selection* rimane inalterato, dando un peso > 0 a tutti i regressori.

Per attuare invece un'operazione di *variable selection* oltre che di *shrinkage*, Tibshirani propose di aggiungere alle stime OLS una penalità in L_1 anziché in L_2 , del tipo :

$$\min_{\beta} \sum_{j=1}^n \left(y_j - \sum_{i=1}^p \beta_i \cdot x_{ij} \right)^2 + \lambda \sum_{i=1}^p |\beta_i|$$

dando origine all'operatore LASSO, che equivale a :

$$\begin{aligned} \min_{\beta} \sum_{j=1}^n \left(y_j - \sum_{i=1}^p \beta_i \cdot x_{ij} \right)^2 \\ \text{sub } \sum_{i=1}^p |\beta_i| \leq t \end{aligned}$$

in cui t è inversamente proporzionale a λ e $\lambda \geq 0$ rappresenta sempre il parametro di penalizzazione che, se posto = 0, fa coincidere le stime a quelle OLS, ma più cresce λ (quindi al tendere di $t \rightarrow 0$) più i parametri del modello si annullano, eliminando quindi le variabili meno significative, in favore delle variabili più importanti per spiegare la variabile Y .

Questa soluzione è definita "sparsa" (perché include alcuni zeri) ed è possibile nella metodologia LASSO e non nella Ridge Regression perché in entrambi i casi il problema di minimizzazione è un problema di ottimo convesso, ma nel secondo caso il vincolo non è differenziabile (per la non differenziabilità del valore assoluto) il che impedisce di ottenere una soluzione in forma chiusa¹⁷ per il minimo globale, come invece avviene con la penalità in L_2 . Per chiarire il concetto è necessario conoscere la definizione e le principali proprietà degli insiemi e delle funzioni convesse¹⁸.

¹⁷ ovvero che può essere risolta in un numero finito di operazioni, si veda "Numerical Solution, Closed-Form Solution" Holton, Glyn (2012)

¹⁸ si veda "Practical Mathematical Optimization" Snyman (2005) e "Numerical Optimization" Nocedal, Wright (2006)

Una forma quadratica del tipo $q(x) = \frac{1}{2}x'Qx + c'x$ è convessa su un insieme convesso se e solo se Q risulta semi-definita positiva. Inoltre, $q(x)$ è strettamente convessa se e solo se Q risulta definita positiva.

La funzione obiettivo dell'algoritmo LASSO è dunque convessa poiché :

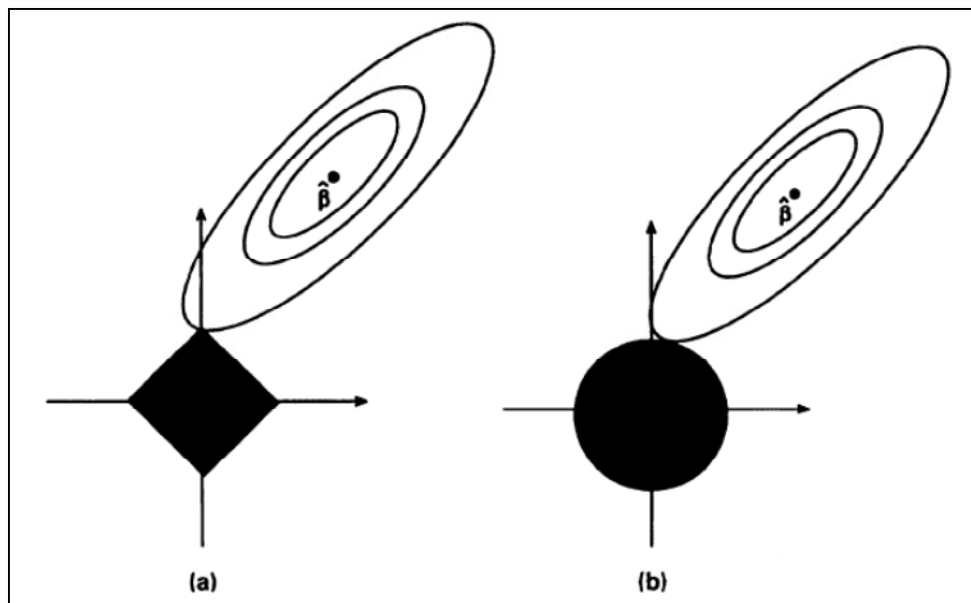
$$\min_{\beta} \sum_{j=1}^n \left(y_j - \sum_{i=1}^p \beta_i \cdot x_{ij} \right)^2 \geq 0$$

equivale alla forma quadratica :

$$(\beta - \hat{\beta}^0)' X'X (\beta - \hat{\beta}^0) \geq 0$$

in cui la matrice $(X'X)$ è semi-definita positiva.¹⁹

Il problema di programmazione quadratica in questione, cerca dunque un ottimo per una funzione convessa, vincolato da regioni anch'esse convesse rappresentate, in uno spazio a due dimensioni, da un cerchio nel caso della Ridge Regression ($\beta_1^2 + \beta_2^2 \leq t$) e da un quadrato nella LASSO ($|\beta_1| + |\beta_2| \leq t$) come rappresentato rispettivamente nei grafici (a) e (b) della seguente figura²⁰, in cui le ellissi rappresentano le isosoluzioni alla funzione obiettivo.



¹⁹ si veda Tibshirani (1996)

²⁰ tratta da Tibshirani (1996)

Si dimostra che nel caso di vincoli rappresentati da regioni convesse, l'ottimo locale corrisponde sempre ad un ottimo globale e tale ottimo si trova sulla frontiera della regione ammissibile, ovvero nel punto in cui le ellissi in figura toccano la regione di vincolo.

Come si nota, nella tecnica LASSO (a) i coefficienti raggiungono il valore nullo perché è possibile imporre un valore di t nel vincolo, tale per cui l'ellisse rappresentante la funzione tocchi la regione quadrata in un suo vertice, proprio nel punto in cui uno dei due coefficienti $\beta_i = 0$. Tale condizione ha probabilità nulla nel caso in cui, invece, la regione di vincolo sia rappresentata da un cerchio come nel caso della Ridge Regression, come si nota nella parte (b) della medesima figura.

2.2 La tecnica LASSO applicata alla matrice di varianze e covarianze

Il metodo LASSO presentato nel precedente paragrafo come stima dei parametri di un modello di regressione lineare, può essere applicato direttamente ad una matrice di varianze e covarianze, risolvendo un problema rilevante nello studio dei modelli grafici gaussiani, in cui si analizza la struttura di una matrice sparsa. Per questo motivo, in questo caso, la tecnica LASSO prende il nome di "graphical lasso" (GLASSO) e permette di effettuare anche un'operazione di *variable selection* tra i titoli, oltre a una penalizzazione come quella proposta dai metodi di *shrinkage* elencati nel capitolo precedente, ai fini di una maggior comprensione della composizione del portafoglio stesso.

In generale, sotto l'ipotesi di normalità p -variata, la log-verosimiglianza campionaria di $y_1, y_2, \dots, y_n \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, posto $\boldsymbol{\mu}=0$, può essere scritta in termini di matrice di precisione $\boldsymbol{\Theta}$, come :

$$\log \det \boldsymbol{\Theta} - tr(\mathbf{S}\boldsymbol{\Theta})$$

dove :

tr indica la traccia,

$\mathbf{S} = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})(y_i - \bar{y})'$ è la matrice di varianze e covarianze campionaria,

$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ è la media campionaria.

La stima di massima verosimiglianza (MLE) di Θ può essere ottenuta massimizzando la seguente equazione per matrici definite positive Θ :

$$\Theta_{MLE} = \max_{\Theta > 0} \{\log \det \Theta - \text{tr}(\mathbf{S}\Theta)\}$$

Come già osservato, la matrice $\mathbf{S}$ e dunque la sua inversa $\mathbf{S}^{-1}$, sono calcolate tramite osservazioni campionarie, che possono portare ad una stima diversa da 0 di alcuni elementi extradiagonali della matrice quando essi potrebbero essere dei veri 0 per l'effettiva incorrelazione tra le variabili in portafoglio.

Per questo motivo uno strumento efficace in molti casi è quello di applicare una funzione di penalizzazione e, quindi, di *variable selection* alla massimizzazione precedente, forzando a zero alcuni elementi in Θ stimati con valori particolarmente bassi.

Dato il campione $y_1, y_2, \dots, y_n$ la funzione di log-verosimiglianza penalizzata tramite la metrica in L_1 può essere scritta come :

$$\log \det \Theta - \text{tr}(\mathbf{S}\Theta) - \lambda \|\Theta\|_1$$

in cui $\|\Theta\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^n |\theta_{ij}|$ è la norma di Θ in L_1 e λ è un parametro di *regolazione* che controlla la dimensione della penalizzazione e dipende sia dai dati che dal numero di elementi non nulli in Θ .

Il problema diventa dunque quello di massimizzare la funzione di log-verosimiglianza penalizzata, su matrici definite positive :

$$\hat{\Theta} = \max_{\Theta > 0} \{\log \det \Theta - \text{tr}(\mathbf{S}\Theta) - \lambda \|\Theta\|_1\}$$

Nel caso in cui $\lambda=0$ la stima coincide con la stima di massima verosimiglianza, mentre per $\lambda>0$ la stima di Θ risulterà sparsa a causa della penalità negli elementi di Θ .

Con la massimizzazione penalizzata, l'obiettivo è quello di trovare un *trade-off*, ovvero una soluzione che nonostante sia sparsa, spieghi adeguatamente i dati : un alto valore di λ corrisponde ad una matrice molto sparsa, mentre un valore basso di λ corrisponde ad una soluzione che si adatta bene ai dati, ma che risulta meno sparsa e quindi meno in grado di effettuare una *variable selection* tra i titoli.

Banerjee et al. (2008) mostrano che il problema di massimizzazione è convesso e considerano la norma di Θ come :

$$\|\Theta\|_1 = \max_{\|U\|_\infty \leq 1} \text{tr}(\Theta U)$$

dove U è una matrice simmetrica con elementi compresi tra -1 ed 1 tale che :

$$\|U\|_\infty = \max_{i=1,\dots,n} \sum_{j=1}^n |u_{ij}| \leq 1$$

La norma di Θ può dunque essere espressa come il massimo della traccia di (ΘU) .

Di conseguenza :

$$-\lambda \|\Theta\|_1 = - \max_{\|U\|_\infty \leq \lambda} \text{tr}(\Theta U) = \min_{\|U\|_\infty \leq \lambda} (-\text{tr}(\Theta U))$$

ed è possibile riscrivere :

$$\begin{aligned} \hat{\Theta} &= \max_{\Theta > 0} \{ \log \det \Theta - \text{tr}(\mathbf{S}\Theta) + \min_{\|U\|_\infty \leq \lambda} (-\text{tr}(\Theta U)) \} = \\ &= \max_{\Theta > 0} \min_{\|U\|_\infty \leq \lambda} \{ \log \det \Theta - \text{tr}(\mathbf{S}\Theta) - \text{tr}(\Theta U) \} = \\ &= \max_{\Theta > 0} \min_{\|U\|_\infty \leq \lambda} \{ \log \det \Theta - \text{tr}[\Theta(\mathbf{S} + \mathbf{U})] \}. \end{aligned}$$

Osservando l'ultima riga dell'equazione, notiamo che alla matrice $\mathbf{S}$, si aggiunge la matrice di perturbazione $\mathbf{U}$. Ciò corrisponde a cercare una stima che massimizzi la log-verosimiglianza nel caso peggiore, cioè su tutte le possibili perturbazioni addizionali alla matrice $\mathbf{S}$.

Ponendo ora $\Theta = (\mathbf{S} + \mathbf{U})^{-1}$ con $(\mathbf{S} + \mathbf{U}) > 0$ definita non-negativa²¹, e ricordando che :

$$\log \det \Theta = - \log \det \Theta^{-1}$$

il problema di ottimizzazione può essere ottenuto come :

$$\min_{\|U\|_\infty \leq \lambda} \{ - \log \det(\mathbf{S} + \mathbf{U}) - p \}$$

che può essere risolto scambiando il minimo con il massimo.

²¹ come sarà dimostrato nella pagina seguente tramite l'utilizzo dei complementi di Schur.

Si ponga infatti $\mathbf{W} = \mathbf{S} + \mathbf{U} = \mathbf{\Theta}^{-1}$

Il problema duale di massima verosimiglianza penalizzata può quindi essere espresso a questo punto in funzione di $\mathbf{\Sigma}$ ed è pari a :

$$\hat{\mathbf{\Sigma}} = \max \{ \log \det \mathbf{W} : \|\mathbf{W} - \mathbf{S}\|_{\infty} \leq \lambda \}$$

il quale stima la matrice di varianze e covarianze, mentre il problema primario stimava la sua inversa. Gli elementi diagonali della soluzione sono pari a : $\Sigma_{ii} = S_{ii} + \lambda$.

Per risolvere il problema duale, è possibile applicare un algoritmo iterativo, in cui il problema di programmazione quadratica è risolto ad ogni step in modo tale da generare una serie di valori via via decrescente, fino a rendere nulli alcuni valori extradiagonali di $\mathbf{\Sigma}$.

Si parta da $\mathbf{W} = \mathbf{S} + \lambda \mathbf{I}$.

Si partizioni la matrice $\mathbf{W}$ come segue :

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_{\setminus p \setminus p} & \mathbf{w}_p \\ \mathbf{w}_p' & W_{pp} \end{bmatrix}$$

dove :

- $\mathbf{W}_{\setminus p \setminus p}$ è la matrice ottenuta rimuovendo la colonna p e la riga p dalla matrice $\mathbf{W}$
- W_{pp} è l'elemento in posizione pp della matrice $\mathbf{W}$
- $\mathbf{w}_p$ è la p-esima colonna rimossa in $\mathbf{W}_{\setminus p \setminus p}$
- $\mathbf{w}_p'$ è la p-esima riga rimossa in $\mathbf{W}_{\setminus p \setminus p}$

Partizionando la matrice $\mathbf{S}$ in modo analogo :

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{\setminus p \setminus p} & \mathbf{s}_p \\ \mathbf{s}_p' & S_{pp} \end{bmatrix}$$

Banerjee et al. (2008) dimostrano che la soluzione corrisponde a :

$$\mathbf{w}_p = \arg \min_{\mathbf{y}} \{ \mathbf{y}' \mathbf{W}_{\setminus p \setminus p}^{-1} \mathbf{y} : \|\mathbf{y} - \mathbf{s}_p\|_{\infty} \leq \lambda \}$$

ovvero un problema di programmazione quadratica vincolato, che si risolve utilizzando una procedura iterativa scambiando di volta in volta le righe e le colonne in modo tale che la

colonna p -esima su cui si agisce sia sempre l'ultima ed aggiornando la stima di $\mathbf{W}$ dopo ogni step, fino a convergenza, ovvero quando :

$$tr(\mathbf{W}^{-1}\mathbf{S}) - p + \lambda \|\mathbf{W}^{-1}\|_1 \leq \varepsilon$$

che si ottiene grazie al fatto che $\mathbf{W} > 0$ ad ogni step.

Utilizzando infatti i complementi di Schur, si ha che :

$$\det(\mathbf{W}) = \det[\mathbf{W}_{\setminus p \setminus p}(W_{pp} - \mathbf{w}_p' \mathbf{W}_{\setminus p \setminus p}^{-1} \mathbf{w}_p)]$$

e, partendo da $\mathbf{W} = \mathbf{S} + \lambda \mathbf{I}$, ad ogni step l'espressione $W_{pp} - \mathbf{w}_p' \mathbf{W}_{\setminus p \setminus p}^{-1} \mathbf{w}_p > 0$

Gli autori²² mostrano che la soluzione del problema duale è equivalente a risolvere il seguente problema di minimi quadrati penalizzato di tipo LASSO :

$$\min_{\boldsymbol{\beta}} \left\{ \frac{1}{2} \|\mathbf{Q}\boldsymbol{\beta} - \mathbf{b}\|_2^2 + \lambda \|\boldsymbol{\beta}\|_1 \right\}$$

in cui : $\mathbf{Q} = \mathbf{W}_{\setminus p \setminus p}^{1/2}$ e $\mathbf{b} = \mathbf{Q}^{-1} \mathbf{s}_p$

Quindi, stimato il vettore di parametri $\boldsymbol{\beta}$ come soluzione del problema LASSO, si ha che $\mathbf{w}_p = \mathbf{W}_{\setminus p \setminus p} \boldsymbol{\beta}$ risolve il problema duale e si procede iterativamente scambiando come detto le righe e le colonne fino a convergenza. Infine $\mathbf{W}^{-1} = \mathbf{\Theta}$.

2.3 Applicazione

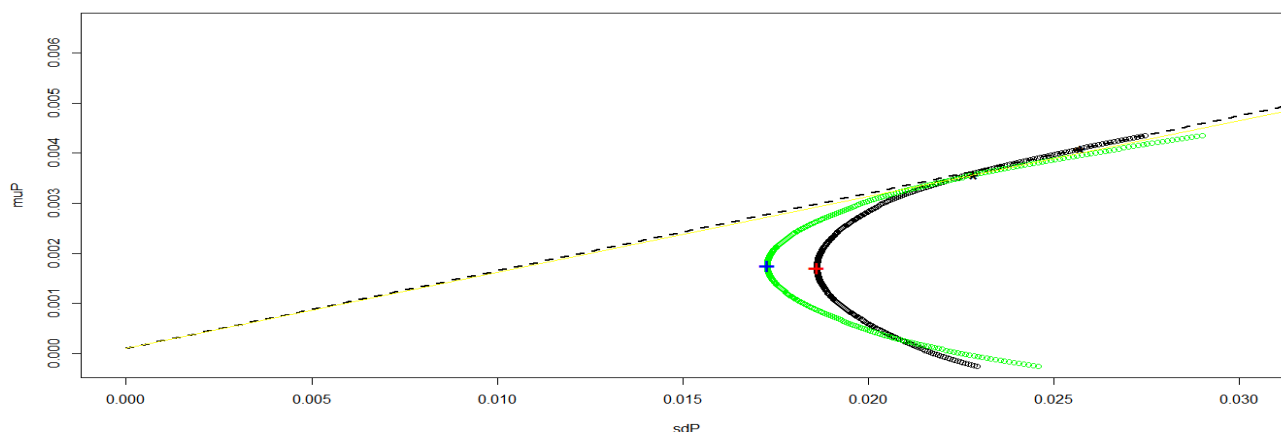
Il software R risolve il problema di minimizzazione vincolata per trovare i portafogli appartenenti alla frontiera efficiente²³, utilizzando la library "quadprog".

Partendo dall'esempio 2 visto nel precedente capitolo, e ipotizzando l'assenza di short sellings per evitare la stima di pesi negativi, è possibile applicare la procedura GLASSO alla matrice di varianze e covarianze dei rendimenti dei 4 titoli azionari tramite i comandi riportati in Appendice.

²² Banerjee (2008)

²³ si veda Ruppert (2011)

Tali comandi producono i seguenti risultati :



PORTAFOGLIO	w1	w2	w3	w4	sdP	MuP
min.var	0.1226	0.2907	0.3127	0.2741	0.0186	0.0017
min.var. CON GLASSO	0.1987	0.2496	0.2847	0.2670	0.0173	0.0018
Tangency	0.0383	0.0833	0.0000	0.8784	0.0257	0.0041
tangency CON GLASSO	0.1490	0.1531	0.0000	0.6979	0.0228	0.0036

L'output riportato visualizza graficamente l'andamento della curva di frontiera efficiente in un grafico E-V sia nel caso in cui la matrice di varianze e covarianze utilizzata sia quella campionaria (in nero), sia nel caso in cui venga applicata la tecnica GLASSO per la sua stima (in verde, con parametro di penalizzazione pari a 0,001).

Il posizionamento del portafoglio tangente e del portafoglio a minima varianza vengono riportati anche in tabella, in cui si mostrano le differenze nella stima dei pesi associati ai 4 titoli azionari sia prima di utilizzare la tecnica GLASSO sia penalizzando la matrice di varianze e covarianze.

Il ruolo del parametro di penalizzazione e l'effetto indotto sui pesi e sulla varianza del portafoglio verranno trattati nel prossimo paragrafo. In questo caso, se si sofferma l'attenzione sul portafoglio a varianza minima (GMVP) l'utilizzo della tecnica porta ad una diminuzione del peso associato ad Unicredit, Fiat e Mediaset a favore dell'unico peso che incrementa, ovvero quello di Intesa San Paolo, producendo così una varianza minore del portafoglio ottimo.

2.4 Parametro di penalizzazione ed effetto su pesi e varianza di portafoglio

Il valore del parametro di penalizzazione λ svolge un ruolo fondamentale nell'algoritmo della tecnica LASSO.

Se esso è posto pari a 0 l'effetto della penalizzazione è nullo, di contro un valore elevato della penalizzazione, per $\lambda \rightarrow \infty$, porta ad una stima della matrice di varianze e covarianze massimamente sparsa, ovvero coincidente alla matrice diagonale, annullando tutti gli elementi extradiagonali.

L'algoritmo GLASSO che è stato presentato nel precedente paragrafo agisce in due possibili modi, in base agli obiettivi dell'utilizzatore : è possibile non penalizzare la diagonale della matrice, ovvero agire solamente sulle covarianze delle variabili portandole verso 0 e mantenendo inalterato il rischio intrinseco di ogni titolo, oppure controbilanciare l'abbattimento delle covarianze aumentando la stima delle varianze campionarie dei titoli azionari, poste sulla diagonale principale.

In base all'opzione scelta e al parametro di penalizzazione λ si ottengono diversi effetti sulle stime dei pesi dei titoli e sulla varianza del portafoglio.

Ponendo l'attenzione da questo momento in avanti sul portafoglio a varianza minima (GMVP), poiché è l'unico portafoglio per il quale non è necessario stimare la frontiera efficiente ed è l'unico portafoglio che dipende solo ed esclusivamente dagli elementi della matrice di varianze e covarianze, si ricorda che²⁴ :

$$\omega_P (min.) = \frac{\Sigma^{-1} \mathbf{1}}{\mathbf{1}' \Sigma^{-1} \mathbf{1}} \quad \sigma_P^2 (min.) = \frac{1}{\mathbf{1}' \Sigma^{-1} \mathbf{1}}$$

Mentre, in generale, la stima della varianza di un portafoglio è data da :

$$\sigma_P^2 = \sum_{i=1}^N \omega_i \sigma_i^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \omega_i \omega_j \sigma_{ij}$$

²⁴ si veda par. 1.7

Si ipotizzi un portafoglio composto da soli 2 titoli azionari.

La matrice di varianze e covarianze Σ è data da :

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

mentre la sua inversa $\Sigma^{-1} = \Theta$ risulta essere :

$$\Sigma^{-1} = \Theta = \begin{bmatrix} \frac{\sigma_2^2}{\det(\Sigma)} & \frac{\sigma_{12}}{\det(\Sigma)} \\ \frac{\sigma_{12}}{\det(\Sigma)} & \frac{\sigma_1^2}{\det(\Sigma)} \end{bmatrix} = \begin{bmatrix} \frac{\sigma_2^2}{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2} & \frac{\sigma_{12}}{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2} \\ \frac{\sigma_{12}}{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2} & \frac{\sigma_1^2}{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2} \end{bmatrix}$$

I risultati ottenuti applicando la tecnica GLASSO con parametro di penalizzazione $\lambda \leq |\sigma_{12}|$ dipendono sia dal segno della covarianza σ_{12} sia dal fatto che si sia scelto di “penalizzare” anche la diagonale principale, oppure si mantengano invariate le varianze σ_i^2 .

Si analizzano ora tutte le possibilità, per comprendere l’effetto della GLASSO sui pesi e sulla varianza di portafoglio caso per caso (si ipotizzi nel seguito $\sigma_1^2 \geq \sigma_2^2$) :

– $\sigma_{12} > 0$ con diagonale penalizzata

(prima di applicare la GLASSO) $\sigma_{P(ante)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2 \omega_1 \omega_2 \sigma_{12}$

(dopo aver applicato la GLASSO) $\sigma_{P(post)}^2 = \omega_1^2 (\sigma_1^2 + \lambda) + \omega_2^2 (\sigma_2^2 + \lambda) + 2 \omega_1 \omega_2 (\sigma_{12} - \lambda)$

Il *gap* di varianza provocato dall’applicazione della procedura GLASSO è quindi dato da :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = \omega_1^2 \lambda + \omega_2^2 \lambda - 2 \omega_1 \omega_2 \lambda = \lambda (\omega_1 - \omega_2)^2$$

Ciò, però, avviene solamente nel caso in cui i pesi ω_1 e ω_2 prima e dopo l’applicazione della GLASSO rimanessero invariati, ovvero quando i titoli presentano lo stesso rischio e quindi $\sigma_1^2 = \sigma_2^2$ oppure quando, indipendentemente dal valore delle varianze, si è nel caso di portafogli equidistribuiti, quindi $\omega_1 = \omega_2 = 0.5$ sia prima che dopo la procedura GLASSO.

In tal caso : $\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = \lambda(\omega_1 - \omega_2)^2 = 0$

Ponendo l'attenzione sul portafoglio GMPV, nel caso generico in cui $\sigma_1^2 > \sigma_2^2$ si ha che :

$$\omega_P = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}} = \begin{cases} \omega_{1(ante)} = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2} \cdot \frac{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} \\ \omega_{2(ante)} = \frac{\sigma_1^2 - \sigma_{12}}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2} \cdot \frac{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} = \frac{\sigma_1^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} \end{cases}$$

mentre dopo la procedura GLASSO :

$$\begin{cases} \omega_{1(post)} = \frac{(\sigma_2^2 + \lambda) - (\sigma_{12} - \lambda)}{(\sigma_1^2 + \lambda) + (\sigma_2^2 + \lambda) - 2(\sigma_{12} - \lambda)} \\ \omega_{2(post)} = \frac{(\sigma_1^2 + \lambda) - (\sigma_{12} - \lambda)}{(\sigma_1^2 + \lambda) + (\sigma_2^2 + \lambda) - 2(\sigma_{12} - \lambda)} \end{cases}$$

Come si nota, essendo $\sigma_1^2 > \sigma_2^2$, si ha che in entrambi i casi $\omega_1 < \omega_2$ cioè il portafoglio ottimo dà un peso minore al titolo con varianza maggiore, a favore di quello con varianza minore.

Per comprendere meglio l'effetto della procedura GLASSO sul portafoglio, è possibile ricavare dalle precedenti formule il *gap* provocato ai pesi del portafoglio prima e dopo l'applicazione della GLASSO :

$$\begin{aligned} \omega_{1(post)} - \omega_{1(ante)} &= \\ &= \frac{(\sigma_2^2 + \lambda) - (\sigma_{12} - \lambda)}{(\sigma_1^2 + \lambda) + (\sigma_2^2 + \lambda) - 2(\sigma_{12} - \lambda)} - \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} = \\ &= \frac{(\sigma_2^2 - \sigma_{12} + 2\lambda)(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}) - (\sigma_2^2 - \sigma_{12})(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12} + 4\lambda)}{(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12} + 4\lambda)(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12})} = \\ &= \dots = \frac{2\lambda(\sigma_1^2 - \sigma_2^2)}{(\sigma_1^2 + \sigma_2^2)^2 + 4\lambda(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}) - 4\sigma_{12}(\sigma_1^2 + \sigma_2^2 - \sigma_{12})} \end{aligned}$$

che è una quantità certamente positiva (nulla solo se $\lambda = 0$) il che significa che il peso minore ω_1 , associato alla varianza maggiore σ_1^2 , aumenta della stessa quantità di cui diminuisce ovviamente ω_2 , portando ad un aumento della varianza di portafoglio.

– $\sigma_{12} < 0$ con diagonale penalizzata

(prima di applicare la GLASSO) $\sigma_{P(ante)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2 \omega_1 \omega_2 \sigma_{12}$

(dopo aver applicato la GLASSO) $\sigma_{P(post)}^2 = \omega_1^2 (\sigma_1^2 + \lambda) + \omega_2^2 (\sigma_2^2 + \lambda) + 2 \omega_1 \omega_2 (\sigma_{12} + \lambda)$

Il *gap* di varianza provocato dall'applicazione della procedura è quindi dato da :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = \omega_1^2 \lambda + \omega_2^2 \lambda + 2 \omega_1 \omega_2 \lambda = \lambda (\omega_1 + \omega_2)^2 = \lambda \geq 0$$

Quindi nel caso in cui i pesi ω_1 e ω_2 prima e dopo l'applicazione della GLASSO rimanessero invariati, ovvero quando $\sigma_1^2 = \sigma_2^2$ oppure si è nel caso di portafogli equidistribuiti, la varianza di portafoglio aumenta proporzionalmente a λ

Ponendo l'attenzione sul portafoglio GMPV, nel caso in cui $\sigma_1^2 > \sigma_2^2$ si ha che :

$$\begin{cases} \omega_{1(ante)} = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} \\ \omega_{2(ante)} = \frac{\sigma_1^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} \end{cases}$$

mentre dopo aver applicato la procedura GLASSO si ha :

$$\begin{cases} \omega_{1(post)} = \frac{(\sigma_2^2 + \lambda) - (\sigma_{12} + \lambda)}{(\sigma_1^2 + \lambda) + (\sigma_2^2 + \lambda) - 2(\sigma_{12} + \lambda)} \\ \omega_{2(post)} = \frac{(\sigma_1^2 + \lambda) - (\sigma_{12} + \lambda)}{(\sigma_1^2 + \lambda) + (\sigma_2^2 + \lambda) - 2(\sigma_{12} + \lambda)} \end{cases}$$

Quindi il *gap* tra i pesi in questo caso è pari a :

$$\begin{aligned} \omega_{1(post)} - \omega_{1(ante)} &= \\ &= \frac{(\sigma_2^2 + \lambda) - (\sigma_{12} + \lambda)}{(\sigma_1^2 + \lambda) + (\sigma_2^2 + \lambda) - 2(\sigma_{12} + \lambda)} - \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} = \\ &= \frac{\sigma_2^2 + \lambda - \sigma_{12} - \lambda}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} - \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} = \end{aligned}$$

$$= \frac{\sigma_2^2 - \sigma_{12} - \sigma_2^2 + \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} = 0$$

In questo caso si nota che i pesi ω_1 e ω_2 non si modificano prima e dopo l'applicazione della GLASSO, quindi : $\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = \lambda(\omega_1 + \omega_2)^2 = \lambda \geq 0$

– $\sigma_{12} > 0$ con diagonale invariata

(prima di applicare la GLASSO) $\sigma_{P(ante)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2 \omega_1 \omega_2 \sigma_{12}$

(dopo aver applicato la GLASSO) $\sigma_{P(post)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2 \omega_1 \omega_2 (\sigma_{12} - \lambda)$

Il *gap* di varianza provocato dall'applicazione della procedura GLASSO è dato da :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = -2 \lambda (\omega_1 \omega_2)$$

Quindi se i pesi ω_1 e ω_2 prima e dopo l'applicazione della GLASSO rimanessero invariati, ovvero nel caso di equidistribuzione, la varianza del portafoglio diminuisce, proporzionalmente al parametro λ .

Ponendo l'attenzione sul portafoglio GMPV, posto sempre $\sigma_1^2 > \sigma_2^2$ si ha che :

$$\begin{cases} \omega_{1(ante)} = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} \\ \omega_{2(ante)} = \frac{\sigma_1^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} \end{cases}$$

$$\begin{cases} \omega_{1(post)} = \frac{\sigma_2^2 - (\sigma_{12} - \lambda)}{\sigma_1^2 + \sigma_2^2 - 2(\sigma_{12} - \lambda)} \\ \omega_{2(post)} = \frac{\sigma_1^2 - (\sigma_{12} - \lambda)}{\sigma_1^2 + \sigma_2^2 - 2(\sigma_{12} - \lambda)} \end{cases}$$

Il *gap* tra i pesi in questo caso è pari a :

$$\begin{aligned} \omega_{1(post)} - \omega_{1(ante)} &= \\ &= \frac{\sigma_2^2 - \sigma_{12} + \lambda}{\sigma_1^2 + \sigma_2^2 - 2(\sigma_{12} - \lambda)} - \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} = \end{aligned}$$

$$\begin{aligned}
&= \frac{(\sigma_2^2 - \sigma_{12} + \lambda)(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}) - (\sigma_2^2 - \sigma_{12})(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12} + 2\lambda)}{(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12} + 2\lambda)(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12})} = \\
&= \dots = \frac{\lambda (\sigma_1^2 - \sigma_2^2)}{(\sigma_1^2 + \sigma_2^2)^2 + 2\lambda(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}) - 4\sigma_{12}(\sigma_1^2 + \sigma_2^2 - \sigma_{12})}
\end{aligned}$$

Come si nota è una quantità certamente positiva (nulla solo se $\lambda = 0$) ma minore rispetto al caso in cui la diagonale venga coinvolta nella penalizzazione. Quindi nonostante il peso associato alla varianza minore aumenti, la varianza del portafoglio diminuisce poiché la diminuzione degli elementi extradiagonali non è controbilanciata come nel caso in cui si coinvolga la diagonale principale nella penalizzazione.

– $\sigma_{12} < 0$ con diagonale invariata

$$(\text{prima di applicare la GLASSO}) \quad \sigma_{P(\text{ante})}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2 \omega_1 \omega_2 \sigma_{12}$$

$$(\text{dopo aver applicato la GLASSO}) \quad \sigma_{P(\text{post})}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2 \omega_1 \omega_2 (\sigma_{12} + \lambda)$$

Il *gap* di varianza provocato dall'applicazione della procedura GLASSO è dato da :

$$\sigma_{P(\text{post})}^2 - \sigma_{P(\text{ante})}^2 = 2 \lambda (\omega_1 \omega_2)$$

Quindi se i pesi ω_1 e ω_2 prima e dopo l'applicazione della GLASSO rimanessero invariati, ovvero nel caso di equidistribuzione, la varianza del portafoglio aumenta, proporzionalmente al parametro λ .

Ponendo l'attenzione sul portafoglio GMPV, posto $\sigma_1^2 > \sigma_2^2$ si ha che :

$$\begin{cases} \omega_{1(\text{ante})} = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} \\ \omega_{2(\text{ante})} = \frac{\sigma_1^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} \end{cases}$$

$$\begin{cases} \omega_{1(\text{post})} = \frac{\sigma_2^2 - (\sigma_{12} + \lambda)}{\sigma_1^2 + \sigma_2^2 - 2(\sigma_{12} + \lambda)} \\ \omega_{2(\text{post})} = \frac{\sigma_1^2 - (\sigma_{12} + \lambda)}{\sigma_1^2 + \sigma_2^2 - 2(\sigma_{12} + \lambda)} \end{cases}$$

Il *gap* tra i pesi è pari a :

$$\begin{aligned}
 \omega_{1(post)} - \omega_{1(ante)} &= \\
 &= \frac{\sigma_2^2 - \sigma_{12} - \lambda}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12} - 2\lambda} - \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} = \\
 &= \frac{(\sigma_2^2 - \sigma_{12} - \lambda)(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}) - (\sigma_2^2 - \sigma_{12})(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12} - 2\lambda)}{(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12} - 2\lambda)(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12})} = \\
 &= \dots = \frac{\lambda (\sigma_2^2 - \sigma_1^2)}{(\sigma_1^2 + \sigma_2^2)^2 - 2\lambda(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}) - 4\sigma_{12}(\sigma_1^2 + \sigma_2^2 - \sigma_{12})}
 \end{aligned}$$

che rappresenta una quantità certamente negativa (nulla solo se $\lambda = 0$) quindi il peso (minore) associato alla varianza maggiore diminuisce, ma la varianza del portafoglio aumenta a causa dell'aumento degli elementi extradiagonali di Σ .

I risultati mostrati nel caso di un portafoglio composto solamente da due titoli azionari, possono essere estesi al caso di portafoglio con più di due componenti.

Si ipotizzi un portafoglio composto da 3 titoli azionari, in una situazione di equiripartizione delle risorse, con pesi $\omega_1 = \omega_2 = \omega_3 = \frac{1}{3}$.

Anche in questo caso è necessario distinguere varie possibilità in base alla scelta di coinvolgere o meno la diagonale principale nella penalizzazione e in base al valore delle covarianze tra i titoli, che possono essere tutte positive, tutte negative, oppure alcune positive ed altre negative.

– $\sigma_{ij} > 0 \quad \forall i,j$ con diagonale penalizzata

Nel caso in cui la tecnica GLASSO viene applicata alla matrice di varianze e covarianze

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix}$$

con covarianze tutte positive e si decida di coinvolgere nella penalizzazione anche la diagonale principale, si ha che :

$$\sigma_{P(ante)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + \omega_3^2 \sigma_3^2 + 2 [\omega_1 \omega_2 \sigma_{12} + \omega_1 \omega_3 \sigma_{13} + \omega_2 \omega_3 \sigma_{23}]$$

$$\begin{aligned} \sigma_{P(post)}^2 &= \omega_1^2 (\sigma_1^2 + \lambda) + \omega_2^2 (\sigma_2^2 + \lambda) + \omega_3^2 (\sigma_3^2 + \lambda) \\ &\quad + 2 [\omega_1 \omega_2 (\sigma_{12} - \lambda) + \omega_1 \omega_3 (\sigma_{13} - \lambda) + \omega_2 \omega_3 (\sigma_{23} - \lambda)] \end{aligned}$$

Il *gap* di varianza provocato dall'applicazione della procedura GLASSO è quindi dato da :

$$\begin{aligned} \sigma_{P(post)}^2 - \sigma_{P(ante)}^2 &= \omega_1^2 \lambda + \omega_2^2 \lambda + \omega_3^2 \lambda - 2 \omega_1 \omega_2 \lambda - 2 \omega_1 \omega_3 \lambda - 2 \omega_2 \omega_3 \lambda = \\ &= \lambda (\omega_1^2 + \omega_2^2 + \omega_3^2 - 2 \omega_1 \omega_2 - 2 \omega_1 \omega_3 - 2 \omega_2 \omega_3) \end{aligned}$$

che, nel caso di equiripartizione, equivale a :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = \lambda \left(\frac{1}{9} + \frac{1}{9} + \frac{1}{9} - \frac{2}{9} - \frac{2}{9} - \frac{2}{9} \right) = -\frac{1}{3} \lambda$$

Quindi, nel caso di più di due titoli azionari, se le covarianze tra i titoli risultano essere tutte positive, l'applicazione della procedura GLASSO porta ad una diminuzione della varianza del portafoglio. Tale risultato si ottiene anche nel caso in cui si decida di mantenere inalterato il rischio di ogni titolo, non coinvolgendo nella penalizzazione la diagonale principale della matrice Σ . In tal caso, infatti :

– $\sigma_{ij} > 0 \quad \forall i,j$ con diagonale invariata

$$\sigma_{P(ante)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + \omega_3^2 \sigma_3^2 + 2 [\omega_1 \omega_2 \sigma_{12} + \omega_1 \omega_3 \sigma_{13} + \omega_2 \omega_3 \sigma_{23}]$$

$$\sigma_{P(post)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + \omega_3^2 \sigma_3^2 + 2 [\omega_1 \omega_2 (\sigma_{12} - \lambda) + \omega_1 \omega_3 (\sigma_{13} - \lambda) + \omega_2 \omega_3 (\sigma_{23} - \lambda)]$$

Il *gap* di varianza provocato dalla procedura GLASSO è quindi dato da :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = -2 \omega_1 \omega_2 \lambda - 2 \omega_1 \omega_3 \lambda - 2 \omega_2 \omega_3 \lambda = -2 \lambda (\omega_1 \omega_2 + \omega_1 \omega_3 + \omega_2 \omega_3)$$

che risulta essere certamente negativo e, nel caso di equiripartizione, equivale a :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = -2 \lambda \left(\frac{1}{9} + \frac{1}{9} + \frac{1}{9} \right) = -\frac{2}{3} \lambda$$

Quindi come nel caso di due soli titoli azionari, l'applicazione della procedura GLASSO porta ad una diminuzione della varianza del portafoglio se le covarianze tra i titoli risultano

essere tutte positive. Questa diminuzione, nel caso di 3 titoli, risulta essere doppia rispetto al caso in cui la penalizzazione coinvolgesse anche la diagonale principale della matrice Σ .

– $\sigma_{ij} < 0 \quad \forall i, j$ con diagonale penalizzata

Se le covarianze tra i titoli azionari fossero tutte negative, si trarrebbero conclusioni analoghe, ma opposte, ai ragionamenti precedenti. Si consideri infatti :

$$\sigma_{P(ante)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + \omega_3^2 \sigma_3^2 + 2 [\omega_1 \omega_2 \sigma_{12} + \omega_1 \omega_3 \sigma_{13} + \omega_2 \omega_3 \sigma_{23}]$$

$$\begin{aligned} \sigma_{P(post)}^2 &= \omega_1^2 (\sigma_1^2 + \lambda) + \omega_2^2 (\sigma_2^2 + \lambda) + \omega_3^2 (\sigma_3^2 + \lambda) \\ &\quad + 2 [\omega_1 \omega_2 (\sigma_{12} + \lambda) + \omega_1 \omega_3 (\sigma_{13} + \lambda) + \omega_2 \omega_3 (\sigma_{23} + \lambda)] \end{aligned}$$

Il *gap* di varianza provocato dalla procedura GLASSO è quindi dato da :

$$\begin{aligned} \sigma_{P(post)}^2 - \sigma_{P(ante)}^2 &= \omega_1^2 \lambda + \omega_2^2 \lambda + \omega_3^2 \lambda + 2 \omega_1 \omega_2 \lambda + 2 \omega_1 \omega_3 \lambda + 2 \omega_2 \omega_3 \lambda = \\ &= \lambda (\omega_1 + \omega_2 + \omega_3)^2 = \lambda \end{aligned}$$

che risulta essere certamente positivo ed equivale al caso con 2 titoli azionari.

Infine, nel caso in cui non venga coinvolta la diagonale principale :

– $\sigma_{ij} < 0 \quad \forall i, j$ con diagonale invariata

$$\sigma_{P(ante)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + \omega_3^2 \sigma_3^2 + 2 [\omega_1 \omega_2 \sigma_{12} + \omega_1 \omega_3 \sigma_{13} + \omega_2 \omega_3 \sigma_{23}]$$

$$\sigma_{P(post)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + \omega_3^2 \sigma_3^2 + 2 [\omega_1 \omega_2 (\sigma_{12} + \lambda) + \omega_1 \omega_3 (\sigma_{13} + \lambda) + \omega_2 \omega_3 (\sigma_{23} + \lambda)]$$

Il *gap* di varianza provocato dalla procedura è quindi dato da :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = 2 \omega_1 \omega_2 \lambda + 2 \omega_1 \omega_3 \lambda + 2 \omega_2 \omega_3 \lambda = 2 \lambda (\omega_1 \omega_2 + \omega_1 \omega_3 + \omega_2 \omega_3)$$

che risulta essere certamente positivo e, nel caso di equiripartizione, equivale a :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = 2 \lambda \left(\frac{1}{9} + \frac{1}{9} + \frac{1}{9} \right) = \frac{2}{3} \lambda$$

Nel caso di più di due titoli azionari, però, al fine di estendere i ragionamenti mostrati, è necessario formulare ulteriori ipotesi, ovvero quando alcune covarianze presenti in Σ risultano essere positive ed alcune negative :

- 2 covarianze positive ed una negativa, con diagonale penalizzata

Si ipotizzi per esempio una matrice di varianze e covarianze in cui il segno “+” rappresenta le covarianze positive tra i titoli e il segno “–” le covarianze negative :

$$\Sigma = \begin{bmatrix} \sigma_1^2 & +\sigma_{12} & +\sigma_{13} \\ +\sigma_{12} & \sigma_2^2 & -\sigma_{23} \\ +\sigma_{13} & -\sigma_{23} & \sigma_3^2 \end{bmatrix}$$

L'applicazione della metodologia di penalizzazione, con coinvolgimento della diagonale principale, porta a :

$$\sigma_{P(ante)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + \omega_3^2 \sigma_3^2 + 2 [\omega_1 \omega_2 \sigma_{12} + \omega_1 \omega_3 \sigma_{13} + \omega_2 \omega_3 \sigma_{23}]$$

$$\begin{aligned} \sigma_{P(post)}^2 &= \omega_1^2 (\sigma_1^2 + \lambda) + \omega_2^2 (\sigma_2^2 + \lambda) + \omega_3^2 (\sigma_3^2 + \lambda) \\ &\quad + 2 [\omega_1 \omega_2 (\sigma_{12} - \lambda) + \omega_1 \omega_3 (\sigma_{13} - \lambda) + \omega_2 \omega_3 (\sigma_{23} + \lambda)] \end{aligned}$$

Il *gap* di varianza provocato dalla procedura GLASSO è quindi dato da :

$$\begin{aligned} \sigma_{P(post)}^2 - \sigma_{P(ante)}^2 &= \omega_1^2 \lambda + \omega_2^2 \lambda + \omega_3^2 \lambda - 2 \omega_1 \omega_2 \lambda - 2 \omega_1 \omega_3 \lambda + 2 \omega_2 \omega_3 \lambda = \\ &= \lambda (\omega_1^2 + \omega_2^2 + \omega_3^2 - 2 \omega_1 \omega_2 - 2 \omega_1 \omega_3 + 2 \omega_2 \omega_3) = \lambda (\omega_1 - \omega_2 - \omega_3)^2 \end{aligned}$$

che è certamente positivo, quindi la varianza del portafoglio aumenta dopo la procedura e nel caso di equiripartizione, vale :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = \lambda \left(\frac{1}{3} - \frac{1}{3} - \frac{1}{3} \right)^2 = \frac{1}{9} \lambda$$

Se non si fosse coinvolta la diagonale principale nella penalizzazione :

- 2 covarianze positive ed una negativa, con diagonale invariata

$$\sigma_{P(ante)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + \omega_3^2 \sigma_3^2 + 2 [\omega_1 \omega_2 \sigma_{12} + \omega_1 \omega_3 \sigma_{13} + \omega_2 \omega_3 \sigma_{23}]$$

$$\sigma_{P(post)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + \omega_3^2 \sigma_3^2 + 2 [\omega_1 \omega_2 (\sigma_{12} - \lambda) + \omega_1 \omega_3 (\sigma_{13} - \lambda) + \omega_2 \omega_3 (\sigma_{23} + \lambda)]$$

Il *gap* di varianza provocato dalla procedura è quindi dato da :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = -2 \omega_1 \omega_2 \lambda - 2 \omega_1 \omega_3 \lambda + 2 \omega_2 \omega_3 \lambda = 2 \lambda (-\omega_1 \omega_2 - \omega_1 \omega_3 + \omega_2 \omega_3)$$

che nel caso di equiripartizione, equivale a :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = 2\lambda \left(-\frac{1}{9} - \frac{1}{9} + \frac{1}{9} \right) = -\frac{2}{9}\lambda$$

Si ipotizzi ora una matrice di varianze e covarianze in cui :

$$\Sigma = \begin{bmatrix} \sigma_1^2 & -\sigma_{12} & -\sigma_{13} \\ -\sigma_{12} & \sigma_2^2 & +\sigma_{23} \\ -\sigma_{13} & +\sigma_{23} & \sigma_3^2 \end{bmatrix}$$

dove il segno “+” rappresenta le covarianze positive tra i titoli e il segno “-” le covarianze negative.

– 2 covarianze negative ed una positiva, con diagonale penalizzata

L’applicazione della metodologia di penalizzazione, con coinvolgimento della diagonale principale, porta a :

$$\sigma_{P(ante)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + \omega_3^2 \sigma_3^2 + 2 [\omega_1 \omega_2 \sigma_{12} + \omega_1 \omega_3 \sigma_{13} + \omega_2 \omega_3 \sigma_{23}]$$

$$\begin{aligned} \sigma_{P(post)}^2 &= \omega_1^2 (\sigma_1^2 + \lambda) + \omega_2^2 (\sigma_2^2 + \lambda) + \omega_3^2 (\sigma_3^2 + \lambda) \\ &\quad + 2 [\omega_1 \omega_2 (\sigma_{12} + \lambda) + \omega_1 \omega_3 (\sigma_{13} + \lambda) + \omega_2 \omega_3 (\sigma_{23} - \lambda)] \end{aligned}$$

Il *gap* di varianza provocato dalla procedura GLASSO è quindi dato da :

$$\begin{aligned} \sigma_{P(post)}^2 - \sigma_{P(ante)}^2 &= \omega_1^2 \lambda + \omega_2^2 \lambda + \omega_3^2 \lambda + 2 \omega_1 \omega_2 \lambda + 2 \omega_1 \omega_3 \lambda - 2 \omega_2 \omega_3 \lambda = \\ &= \lambda (\omega_1^2 + \omega_2^2 + \omega_3^2 + 2 \omega_1 \omega_2 + 2 \omega_1 \omega_3 - 2 \omega_2 \omega_3) \end{aligned}$$

che è certamente positivo, quindi la varianza del portafoglio aumenta dopo la procedura e nel caso di equiripartizione, vale :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = \lambda \left(\frac{1}{3} + \frac{2}{9} + \frac{2}{9} - \frac{2}{9} \right) = \frac{5}{9}\lambda$$

Se non si fosse coinvolta la diagonale principale nella penalizzazione :

– 2 covarianze negative ed una positiva, con diagonale invariata

$$\sigma_{P(ante)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + \omega_3^2 \sigma_3^2 + 2 [\omega_1 \omega_2 \sigma_{12} + \omega_1 \omega_3 \sigma_{13} + \omega_2 \omega_3 \sigma_{23}]$$

$$\sigma_{P(post)}^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + \omega_3^2 \sigma_3^2 + 2 [\omega_1 \omega_2 (\sigma_{12} + \lambda) + \omega_1 \omega_3 (\sigma_{13} + \lambda) + \omega_2 \omega_3 (\sigma_{23} - \lambda)]$$

Il *gap* di varianza provocato dalla procedura è quindi dato da :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = 2 \omega_1 \omega_2 \lambda + 2 \omega_1 \omega_3 \lambda - 2 \omega_2 \omega_3 \lambda = 2 \lambda (\omega_1 \omega_2 + \omega_1 \omega_3 - \omega_2 \omega_3)$$

che nel caso di equiripartizione, equivale a :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = 2 \lambda \left(\frac{1}{9} + \frac{1}{9} - \frac{1}{9} \right) = \frac{2}{9} \lambda$$

Come si nota, dunque, l'estensione di questi risultati al generico caso di portafoglio con N titoli azionari, è una procedura piuttosto complessa, da valutare caso per caso, ma che può spesso portare a conclusioni univoche. Nella maggior parte dei casi, infatti, i titoli appartenenti al medesimo mercato finanziario tendono ad assumere covarianze positive tra loro, poiché influenzate simultaneamente dall'andamento globale del mercato stesso e, inoltre, un investitore non trova spesso utile coinvolgere nella penalizzazione la diagonale principale, che rappresenta il rischio intrinseco di ogni titolo, da mantenere inalterato.

In questo contesto, dunque, con covarianze positive tra gli N titoli e diagonale principale invariata, si è potuto dimostrare come l'applicazione della tecnica GLASSO comporti la composizione di un portafoglio con varianza minore del portafoglio iniziale, non penalizzato, con una decremento pari a :

$$\sigma_{P(post)}^2 - \sigma_{P(ante)}^2 = -2 \lambda \sum_i \sum_j \omega_i \omega_j$$

CAP. 3

GRAFI

3.1 Cenni alla teoria dei grafi.

La tecnica LASSO presentata nel precedente capitolo, applicata alla matrice di varianze e covarianze di un portafoglio azionario, permette all'investitore una migliore comprensione della composizione del portafoglio stesso e quindi la possibilità di effettuare *portfolio selection*, nell'ottica di un'ottimizzazione delle risorse a propria disposizione.

Il presente capitolo, insieme al successivo, ha come obiettivo principale proprio quello di mostrare gli effetti che l'applicazione della tecnica LASSO produce sull'ottimizzazione di portafoglio e, a tal fine, risulta utile servirsi di uno strumento preso in prestito dalla teoria dei grafi che permetta di evidenziare le dipendenze tra i titoli azionari e la struttura del portafoglio ottimo.

Si considerino n campioni indipendenti tratti da una distribuzione gaussiana p -variata, che rappresentano i rendimenti dei titoli azionari : $y_1, y_2, \dots, y_n \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, dove la matrice di varianze e covarianze $\boldsymbol{\Sigma}$ è oggetto di stima e renderla *sparsa* permetterebbe di catturare maggiormente le dipendenze tra i titoli presenti in portafoglio.

Nell'ambito di un portafoglio azionario, per approfondire il concetto di diversificazione, è l'indipendenza condizionata a giocare un ruolo fondamentale, ovvero l'indipendenza tra due titoli, dati tutti gli altri. Molti autori pongono quindi la loro attenzione sulla sparsità della matrice inversa $\boldsymbol{\Theta}$, piuttosto che di $\boldsymbol{\Sigma}$.

La matrice $\boldsymbol{\Theta}$ ²⁵ ha infatti una specifica interpretazione : mentre la presenza di zeri in una matrice di varianze e covarianze $\boldsymbol{\Sigma}$ indica un'incorrelazione tra le corrispondenti variabili marginali, se l'elemento ij -esimo di $\boldsymbol{\Theta}$, ovvero θ_{ij} , è pari a zero implica che le corrispondenti variabili y_i e y_j sono *condizionalmente* indipendenti, date le altre variabili.

²⁵ nota in letteratura anche con il nome di "matrice di precisione".

Il concetto di indipendenza condizionale è di centrale importanza soprattutto nel caso di modelli gaussiani. Presa infatti una variabile casuale p-dimensionale $Y=(y_1, y_2, \dots, y_p)$ l'inversa della matrice di varianze e covarianze è rappresentata da :

$$\begin{bmatrix} \theta_{11} & \theta_{12} & \dots & \theta_{1p} \\ \theta_{21} & \theta_{22} & \dots & \theta_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{p1} & \theta_{p1} & \dots & \theta_{pp} \end{bmatrix}$$

La distribuzione bivariata di (Y_1, Y_2) date tutte le altre $(Y_3, \dots, Y_p)$ ha covarianza pari a :

$$\begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix}^{-1} = \frac{1}{\theta_{11}\theta_{22} - (\theta_{12})^2} \begin{pmatrix} \theta_{22} & -\theta_{21} \\ -\theta_{12} & \theta_{11} \end{pmatrix}$$

Il coefficiente di correlazione tra Y_1 e Y_2 date tutte le altre variabili è chiamato coefficiente di correlazione parziale e può essere visto in termini di Θ , come :

$$\rho_{12|3,\dots,p} = \frac{-\theta_{12}}{\theta_{11}\theta_{22} - (\theta_{12})^2} \bigg/ \left[\frac{\theta_{22}}{\theta_{11}\theta_{22} - (\theta_{12})^2} \frac{\theta_{11}}{\theta_{11}\theta_{22} - (\theta_{12})^2} \right]^{1/2} = \frac{-\theta_{12}}{(\theta_{11}\theta_{22})^{1/2}}$$

da cui si nota che :

$$\rho_{12|3,\dots,p} = 0 \Leftrightarrow \theta_{12} = 0$$

Nel caso gaussiano, l'incorrelazione parziale implica l'indipendenza e viceversa, quindi due variabili i e j sono condizionalmente indipendenti (date tutte le altre) se e solo se il corrispondente elemento θ_{ij} della matrice di precisione Θ è nullo.

In questi casi, dunque, l'attenzione viene rivolta soprattutto alla sparsità della matrice inversa della matrice di varianze e covarianze, $\Theta=\Sigma^{-1}$, piuttosto che di Σ stessa, poiché, se l'elemento θ_{ij} è pari a zero implica che le corrispondenti variabili y_i e y_j sono condizionalmente indipendenti, date le altre variabili.

In questo contesto, è quindi utile introdurre alcuni concetti chiave della *teoria dei grafi*, accennandone le principali definizioni.

Un grafo G è una coppia $G=(V,E)$ dove V è un insieme finito (non vuoto), detto insieme dei vertici, ed E è detto insieme dei lati (o degli archi). I lati sono dunque coppie non ordinate di vertici, cioè un lato è identificato con i vertici che ne costituiscono le estremità, quindi dati due vertici, ci può essere al più un solo lato che li congiunge.

Dato un grafo G , gli insiemi dei vertici e dei lati di G sono anche indicati con $V(G)$ ed $E(G)$ poiché la lettera E è l'iniziale di "edge", mentre V sta per "vertex".

Sia dunque $V = \{v_1, v_2, \dots, v_n\}$ l'insieme dei vertici di un grafo e sia $[V]^2$ l'insieme di tutti i possibili sottoinsiemi di V di cardinalità 2. Dato che l'insieme E dei lati è un sottoinsieme di $[V]^2$, il numero dei grafi che è possibile costruire con n vertici è dato dalla cardinalità dell'insieme delle parti di $[V]^2$, ovvero : $2^{\binom{n}{2}} = 2^{\frac{n(n-1)}{2}}$.

Due vertici distinti sono *adiacenti* se esiste un lato che li congiunge. Due lati distinti sono *adiacenti* se hanno un vertice in comune.

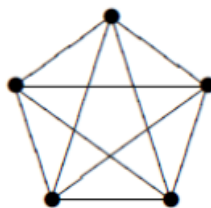
Per rappresentare graficamente un grafo si disegna un punto per ogni vertice e un segmento (o un arco di curva) per ogni lato, in modo che tutte le relazioni di adiacenza siano rispettate. Il modo in cui vengono disposti i punti corrispondenti ai vertici e gli archi corrispondenti ai lati è irrilevante.

L'*ordine* di un grafo G , indicato con $|G|$ oppure con $n(G)$, è dato dal numero dei vertici di G . Il numero dei lati è invece detto *grandezza* di G ed è indicato con $||G||$ oppure con $m(G)$.

Un grafo è detto *completo* se i suoi vertici sono a due a due adiacenti, cioè se ogni coppia di vertici è collegata da un lato.

Un grafo completo con n vertici è indicato con K^n ed ha $\binom{n}{2}$ lati.

Per esempio, per $n=5$ si ha il grafo K^5 :



Un grafo *orientato* (o diretto) è un grafo in cui a ogni lato è stato assegnato un orientamento.

Un grafo può anche essere descritto in termini di matrici. Sia $G = (V, E)$ un grafo. Sia $V = \{v_1, v_2, \dots, v_n\}$ l'insieme dei vertici ed $E = \{e_1, e_2, \dots, e_m\}$ l'insieme dei lati di G . Per ogni coppia di vertici $(v_i, v_j) \in V$, poniamo :

$$a_{ij} = \begin{cases} 1 & \text{se } (v_i, v_j) \in E \\ 0 & \text{se } (v_i, v_j) \notin E \end{cases}$$

La matrice $A(G) = (a_{ij})$ è detta *matrice di adiacenza* del grafo G ed è una matrice quadrata di ordine n con coefficienti 0 o 1, simmetrica se il grafo non è orientato, ovvero se $a_{ij} = a_{ji}$ cioè $(v_i, v_j) = (v_j, v_i)$ come nel caso che verrà trattato nel presente elaborato, in cui il grafo rappresenta la matrice di precisione Θ , con $\theta_{ij} = \theta_{ji}$.

3.2 Applicazione in finanza.

Nel caso di modelli gaussiani, due variabili i e j sono condizionalmente indipendenti, date tutte le altre, se e solo se il corrispondente elemento θ_{ij} della matrice di precisione Θ è nullo.

In ambito finanziario, è possibile dunque applicare il modello grafico gaussiano in cui ogni nodo rappresenta una variabile, ovvero un titolo azionario presente in portafoglio, e due nodi sono collegati tra loro da un lato (o arco) se la corrispondente correlazione parziale non è nulla. Quindi risulta importante la sparsità della matrice di precisione $\Theta = \Sigma^{-1}$.

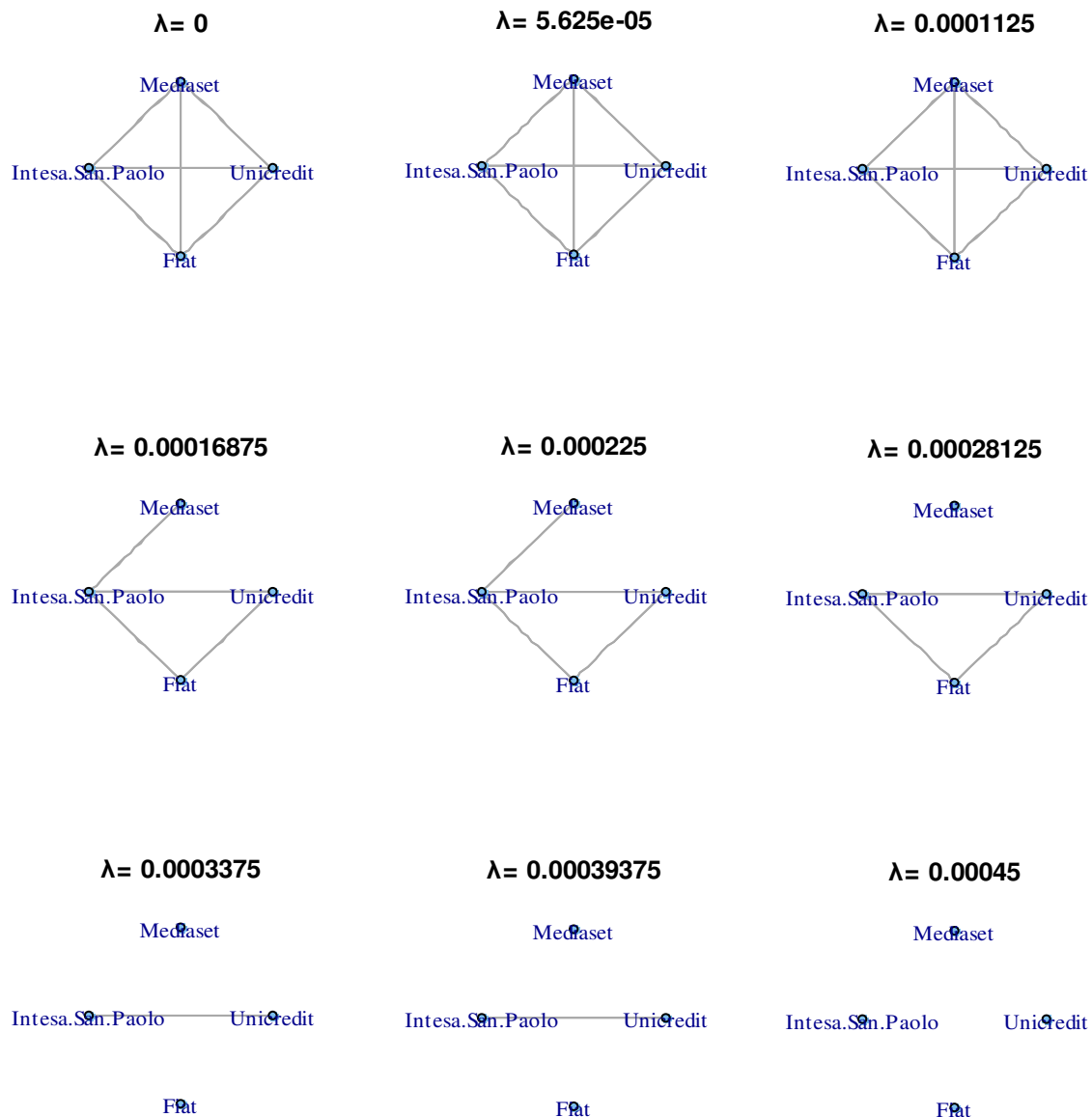
Applicando la tecnica GLASSO, si penalizzano gli elementi extradiagonali della matrice di varianze e covarianze, rendendola così sparsa, ma all'aumentare del parametro di penalizzazione λ assume importanza anche la corrispondente sparsità della matrice inversa di Σ .

Se, infatti, per $\lambda = 0$ tutti i rendimenti dei titoli azionari presenti in portafoglio sono condizionalmente *correlati* tra loro, e quindi la matrice di precisione Θ non contiene zeri e il grafo associato è da considerarsi *completo*, all'aumentare della penalizzazione si rende sparsa la matrice di precisione Θ e il grafo perde tutti quei lati associati alle variabili (titoli) condizionalmente indipendenti tra loro.

Associare dunque un grafo ad un portafoglio azionario, all'aumentare di λ , dà una chiara indicazione sulla selezione del portafoglio stesso : evidenziando quei titoli che a partire da un certo $\lambda > 0$ non risultano più connessi con nessun altro, si mostra come essi apportino un contributo nullo alla diversificazione.

Si applichi ora all'esempio 2 mostrato nei capitoli precedenti, la metodologia supportata dal software R per la costruzione di un grafo non orientato, tramite la library denominata "graph", attraverso i comandi riportati in Appendice.

Tali comandi producono un grafo per ogni livello di penalizzazione scelto :



Dall'output mostrato si nota ciò che si era preliminarmente ipotizzato : con $\lambda = 0$ tutti i titoli sono connessi tra di loro nel grafo; ma all'aumentare della penalizzazione alcuni archi non congiungono più quei titoli condizionalmente indipendenti, che invece sembravano essere correlati tra loro se non si fosse applicata la procedura GLASSO, fino a un certo livello di penalizzazione λ in cui solo i due titoli appartenenti a gruppi bancari sembrano essere dipendenti tra loro, come ci si aspetterebbe nella realtà e come evidenziato nel paragrafo 1.10. Infine, a partire da un livello di penalizzazione $\lambda > \max(\sigma_{ij})$, tutti i titoli risultano essere incorrelati tra loro, ovvero la matrice di precisione risulta essere diagonale e il grafo perde tutti i lati.

CAP. 4

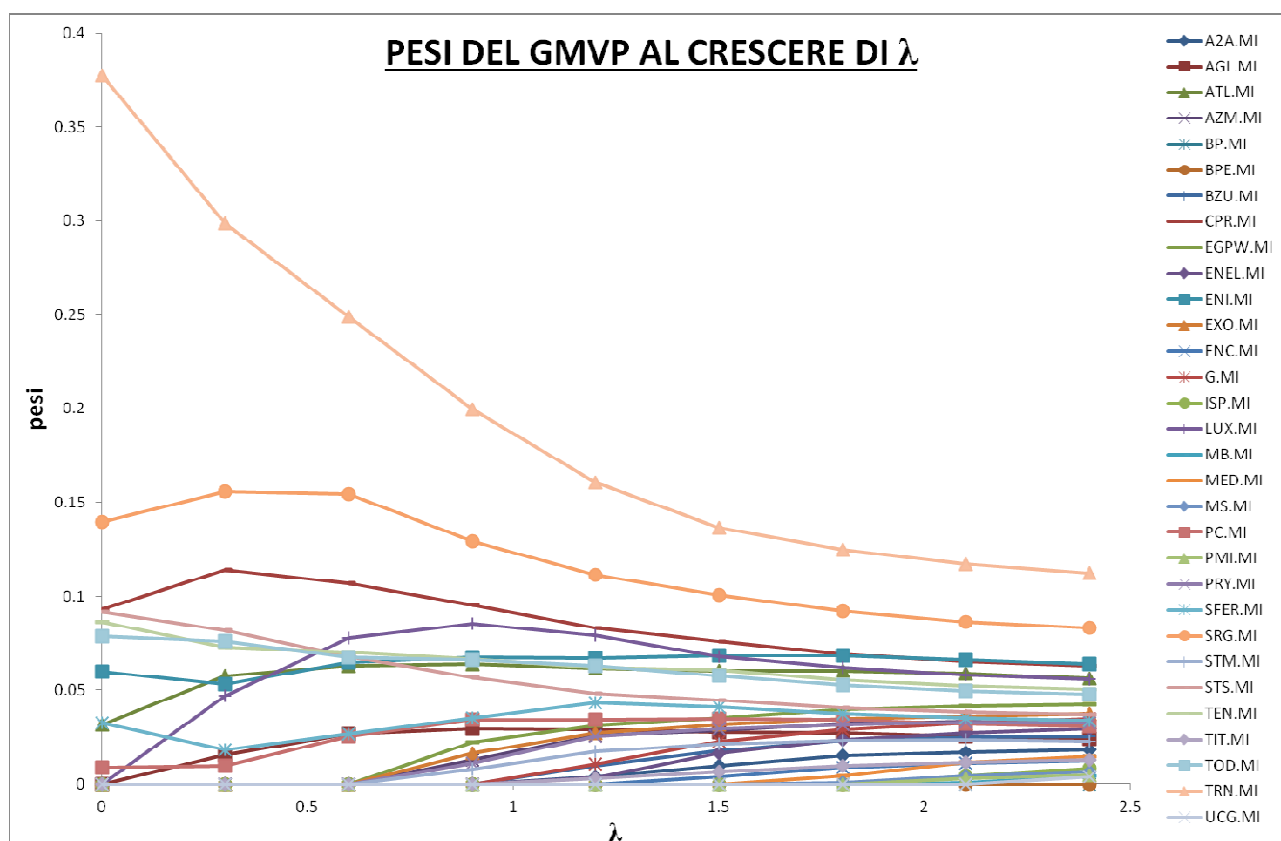
UTILIZZO DELLA TECNICA LASSO IN FINANZA

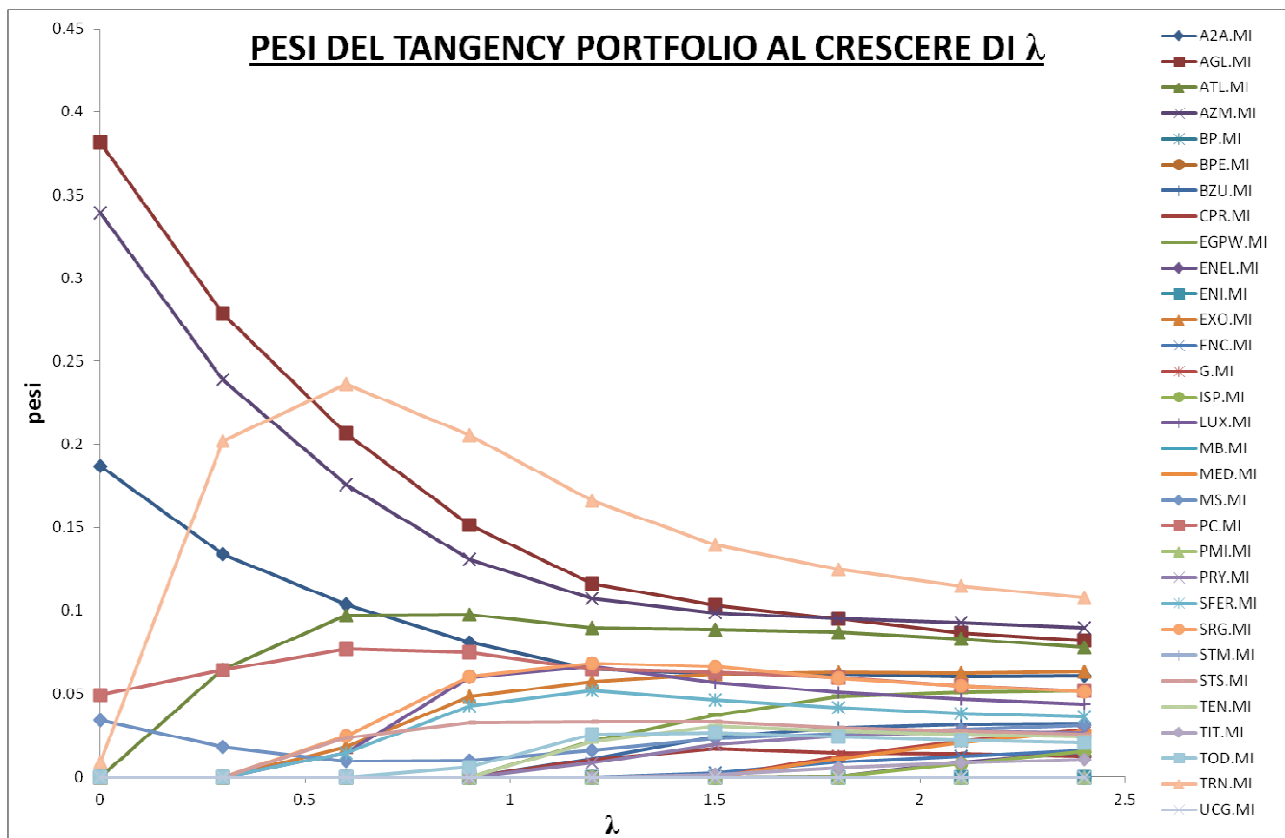
4.1 Composizione del portafoglio ottimo

Nel capitolo precedente si è mostrato come, al variare del parametro di penalizzazione λ , si modifichino le interconnessioni tra i titoli presenti nel portafoglio e nell'intero mercato azionario.

L'utilizzo dell'algoritmo LASSO applicato alla matrice di varianze e covarianze dei rendimenti dei titoli azionari, comporta ulteriori innovazioni nello studio del portafoglio ottimo a disposizione di un investitore.

Se infatti si focalizza l'attenzione sui pesi ω_i del portafoglio a varianza minima (GMVP) e del portafoglio tangente, fissato un arco temporale, si nota come all'aumentare del parametro di penalizzazione λ , si modifica la composizione del portafoglio stesso :





In entrambi i casi rappresentati nelle figure precedenti, si prendano in considerazione i titoli azionari appartenenti all'indice FTSEMIB della Borsa Italiana nell'ultimo anno solare.

Senza applicare alcuna penalizzazione, quindi con parametro $\lambda=0$, la composizione del portafoglio ottimo, che sia esso rappresentato dal GMVP o dal Tangency portfolio, contiene solo alcuni titoli azionari, dando peso nullo a tutti gli altri.

Impiegando una tecnica di penalizzazione di tipo LASSO, con parametro λ via via crescente, si modifica la struttura di dipendenza tra i titoli, ottenendo una diversa composizione del portafoglio ottimo, tale per cui alcuni titoli che inizialmente, senza applicare la LASSO, presentavano pesi maggiori, vedono diminuire il loro apporto alla composizione del portafoglio ottimo, a favore di quei titoli che entrano a far parte del portafoglio, con un peso $\omega_i > 0$, solamente a partire da un certo livello di penalizzazione in poi.

Si noti che raggiunto un determinato parametro di penalizzazione λ (nel caso in figura $\lambda=2.4$) ovvero raggiunta la penalizzazione massima tale per cui tutti gli elementi extradiagonali della matrice di varianze e covarianze risultano nulli, la composizione del portafoglio ottimo tende a stabilizzarsi verso valori proporzionali al reciproco degli elementi

diagonali di Σ . L'avvicinamento tra loro dei pesi di portafoglio non porta comunque tutti i pesi allo stesso valore $\omega_i > 1/p$ (dove p è il numero di titoli presenti in portafoglio) poiché non tutti i titoli hanno lo stesso valore di rischio intrinseco, ovvero la stessa varianza. Tale situazione si otterrebbe infatti solo nel caso in cui la diagonale principale della matrice di varianze e covarianze, cui viene applicata la tecnica LASSO, avesse elementi tutti uguali tra loro (si pensi per esempio all'applicazione dell'algoritmo LASSO alla matrice di correlazione, in luogo delle varianze e covarianze).

Nella realtà dei casi finanziari, dunque, anche un'alta penalizzazione alla matrice Σ produce un portafoglio ottimo in cui viene dato peso minore ai titoli con rischio intrinseco più alto e peso maggiore ai titoli meno rischiosi, al netto delle relazioni tra i titoli che, proprio da un certo valore di λ in poi, vengono annullate.

4.2 Scelta del parametro di penalizzazione λ

Nei capitoli precedenti, è già stato sottolineato, soprattutto nel paragrafo 2.4, quanto la scelta del parametro di penalizzazione λ sia fondamentale nell'algoritmo LASSO applicato in ambito finanziario, sia per quanto riguarda l'effetto che esso produce sui pesi e sulla varianza del portafoglio ottimo, sia per la composizione stessa del portafoglio, messa in luce soprattutto grazie all'utilizzo dei grafi e ai risultati ottenuti nel precedente paragrafo.

Nella letteratura moderna, in materia di stima di massima verosimiglianza penalizzata, vengono proposte diverse tecniche per la scelta del λ ottimo, tra le quali la più diffusa è certamente rappresentata da una metodologia di cross-validation²⁶.

In generale, le tecniche di cross-validation presuppongono di dividere ripetutamente i dati campionari in una parte (*training set*) usata per la stima del modello e in una seconda parte (*validation set*) che tale modello cerca di prevedere, in maniera tale da stimare un errore predittivo confrontando il modello con la realtà dei dati osservati.

²⁶ si veda Friedman, Hastie, Tibshirani (2008)

Nel caso in esame, dunque, avendo a disposizione numerosi dati finanziari, al fine di individuare il λ ottimo, si divide il campione in k sotto-gruppi. Per ogni sotto-gruppo si stima il valore della log-verosimiglianza penalizzata, per ogni valore di λ :

$$\log \det \Theta - \text{tr}(\mathbf{S}\Theta) - \lambda \|\Theta\|_1$$

Lo stessa stima si effettua per la parte restante del campione, formato da $(k - 1)/k$ valori.

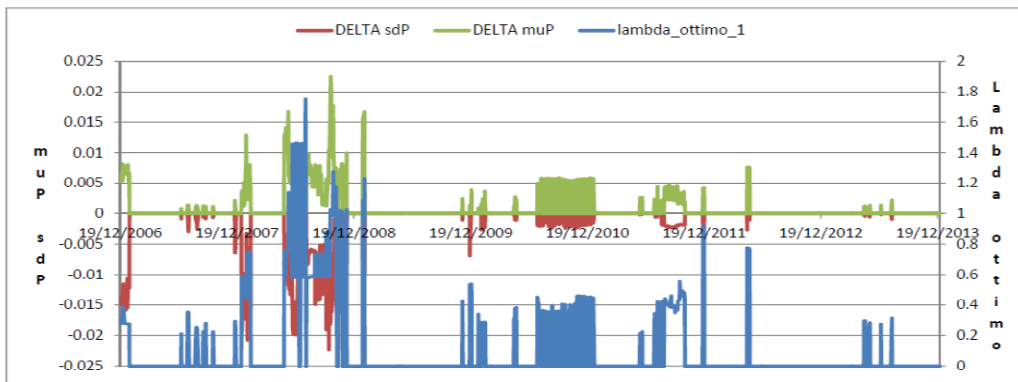
Il parametro λ ottimo da utilizzare per l'algoritmo LASSO è quello che minimizza la distanza (media) tra la log-verosimiglianza penalizzata calcolata sul *validation set* rispetto alla restante parte del campione.

Sempre tramite l'utilizzo della cross-validation, Meinshausen e Bühlmann (2006) propongono una stima ottimale del parametro di penalizzazione λ dell'algoritmo LASSO, proprio in termini di composizione di un grafo. Nonostante la loro applicazione non sia in ambito finanziario, l'obiettivo è quello di stimare un valore di λ risolvendo un problema di *Neighborhood selection*, ovvero analizzando il numero di variabili connesse ad ogni nodo del grafo.

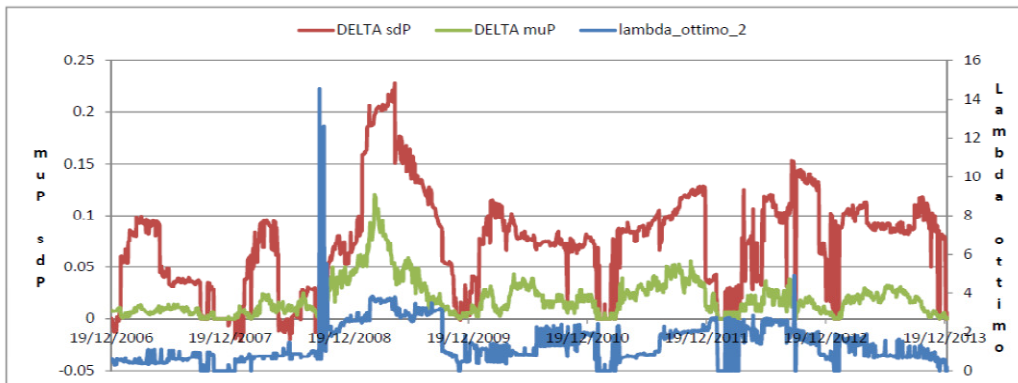
Nel contesto di ottimizzazione di portafoglio preso in considerazione nel presente elaborato, la nostra proposta di scelta del λ ottimo si discosta dalle già note e citate metodologie presenti in letteratura riguardo la cross-validation.

Dato un portafoglio ottimo, come il Global Minimum Variance Portfolio, l'obiettivo è quello di applicare un algoritmo LASSO tale per cui il portafoglio risultante, una volta modificata la struttura di dipendenza tra i titoli, si tramuti in un portafoglio con rendimento atteso maggiore o con una misura di rischio minore (o, addirittura, con entrambi gli aspetti simultaneamente) rispetto al portafoglio originario, con la matrice di varianze e covarianze campionaria senza l'applicazione della LASSO.

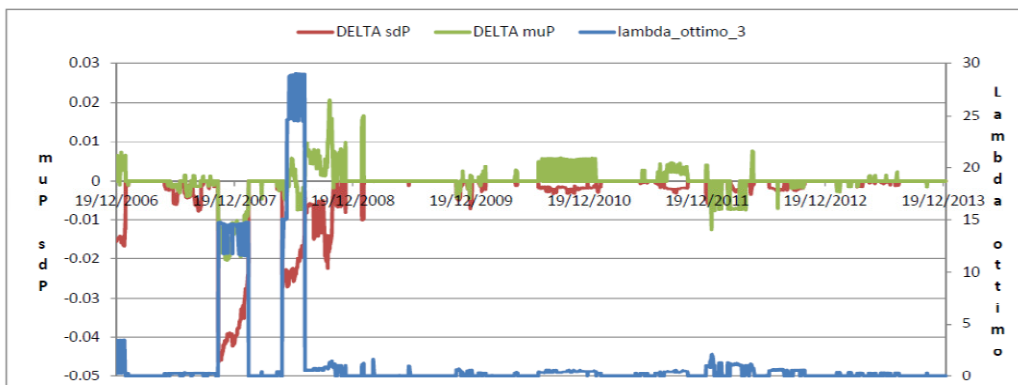
L'idea per la scelta del λ ottimo è dunque quella proposta nelle tre diverse strategie riportate dai seguenti grafici :



GMVP con:
 $\mu_P(\text{lasso}) > \mu_F$
 $\text{sd}_P(\text{lasso}) < \text{sd}_P$



GMVP con:
 $\mu_P(\text{lasso}) > \mu_F$



GMVP con:
 $\text{sd}_P(\text{lasso}) < \text{sd}_P$

Si considerino i rendimenti dei titoli azionari appartenenti all'indice FTSEMIB dall'anno 2006 all'anno 2013. Per ogni grafico si calcoli, per ogni tempo t , il rendimento atteso e lo scarto quadratico medio del Global Minimum Variance Portfolio, con una finestra temporale relativa ai 180 giorni precedenti.

La prima strategia proposta è quella di assumere, per ogni istante temporale t , un λ ottimo diverso da 0 se la procedura di penalizzazione (con λ che varia iterativamente tra 0 e il valore della massima covarianza osservata²⁷) produce un portafoglio con rendimento atteso maggiore e simultaneamente scarto quadratico medio minore del portafoglio GMVP iniziale, non penalizzato.

²⁷ in valore assoluto, poiché quando $\lambda \geq \max(|\sigma_{ij}|)$ la matrice Σ risulta essere diagonale.

Come si vede dal primo grafico, sono solo alcuni i casi in cui l'applicazione della tecnica LASSO produce questa duplice ottimizzazione rispetto al GMVP non penalizzato, adottando un λ ottimo variabile nel tempo, con un massimo di 1.7.

Nella seconda strategia, riportata nel secondo grafico, la proposta è quella di assumere, per ogni istante temporale t , un λ ottimale diverso da 0 se grazie alla procedura di penalizzazione si ottiene un portafoglio con solamente rendimento atteso maggiore del portafoglio GMVP iniziale, al di là dell'eventuale incremento della misura di rischio. Come si nota, in questo secondo caso, essendo il vincolo meno restrittivo, sono molti i periodi in cui la condizione è soddisfatta e quindi, salvo rari periodi particolarmente *outlier*, si ottiene un λ ottimo che oscilla tra 0 e 3.

Nel terzo caso, il valore del λ ottimo viene identificato se, per ogni tempo t , la tecnica LASSO produce un Global Minimum Variance Portafoglio con scarto quadratico medio minore rispetto al portafoglio iniziale. Ciò avviene quando nel terzo ed ultimo grafico la linea rossa, che identifica la differenza tra $\sigma_{GMVP}(post - LASSO) - \sigma_{GMVP}(ante - LASSO)$ risulta essere minore di 0, ovvero sotto l'asse delle ascisse, indipendentemente dal valore del rendimento atteso del portafoglio, il cui "gap" tra dopo la LASSO e prima della LASSO è identificato dalla linea verde.

La linea blu, con scala riportata sull'asse delle ordinate a destra nei tre grafici, rappresenta invece il valore del λ ottimo per ogni tempo t .

4.3 Dinamiche del mercato azionario

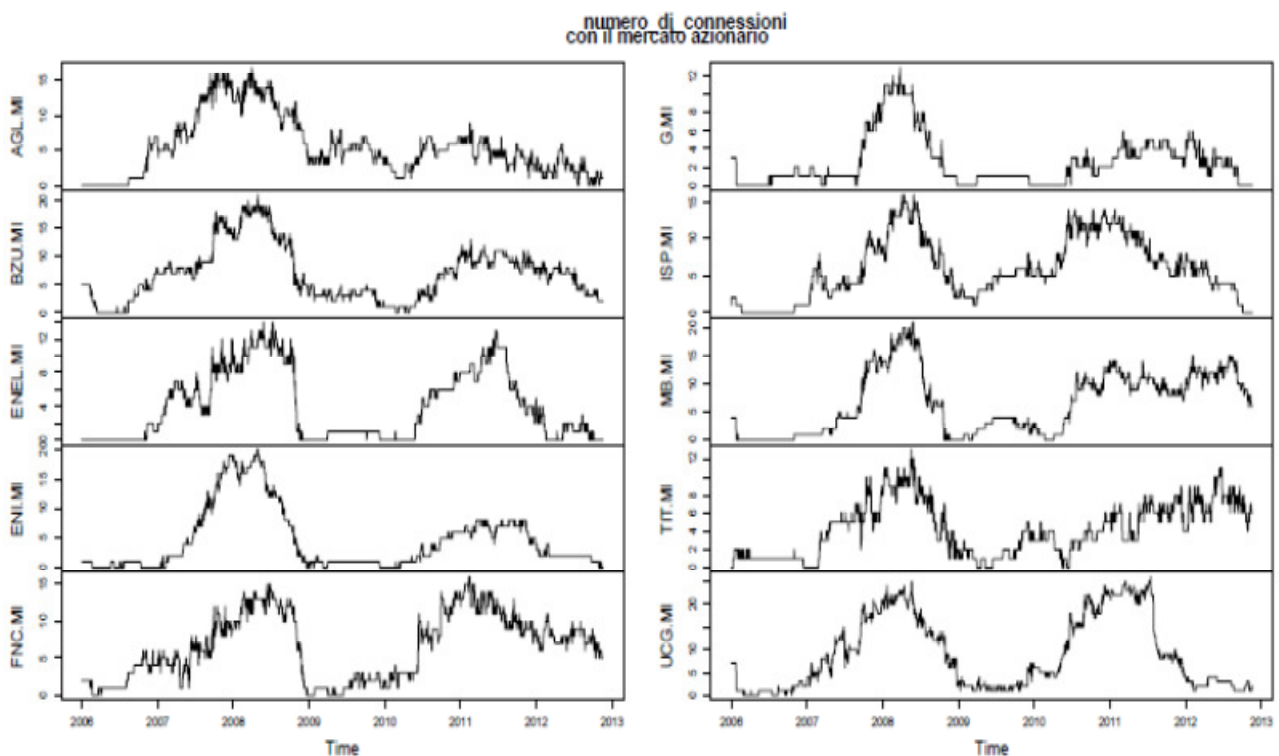
Fino a questo momento, lo studio del portafoglio ottimo con l'applicazione dell'algoritmo LASSO è sempre stato effettuato in un'ottica statica, ovvero fissato un istante temporale t .

In quest' ultima parte dell'elaborato invece, si mostra come l'analisi può essere svolta in un contesto dinamico, in cui al variare del tempo, nuove osservazioni riguardanti il mercato azionario influenzano la composizione del portafoglio : i pesi del portafoglio e le relazioni di dipendenza tra i titoli subiscono quindi cambiamenti dovuti alle nuove situazioni di mercato.

Si ponga l'attenzione al numero di connessioni che ogni titolo ha con i restanti titoli presenti nel portafoglio o, a maggior ragione, con l'intero mercato azionario. Si è già visto attraverso la composizione dei grafi, come all'aumentare del parametro di penalizzazione λ , il numero di connessioni attive tra i titoli (ovvero il numero di correlazioni diverse da 0) diminuisce, in qualsiasi istante temporale t .

Fissato un livello di penalizzazione $\lambda > 0$ con una delle tecniche illustrate nel paragrafo precedente, è possibile studiare come variano le interconnessioni tra i titoli azionari nel tempo e che relazione c'è tra di esse e l'andamento generale del mercato azionario, rappresentato dall'indice FTSEMIB.

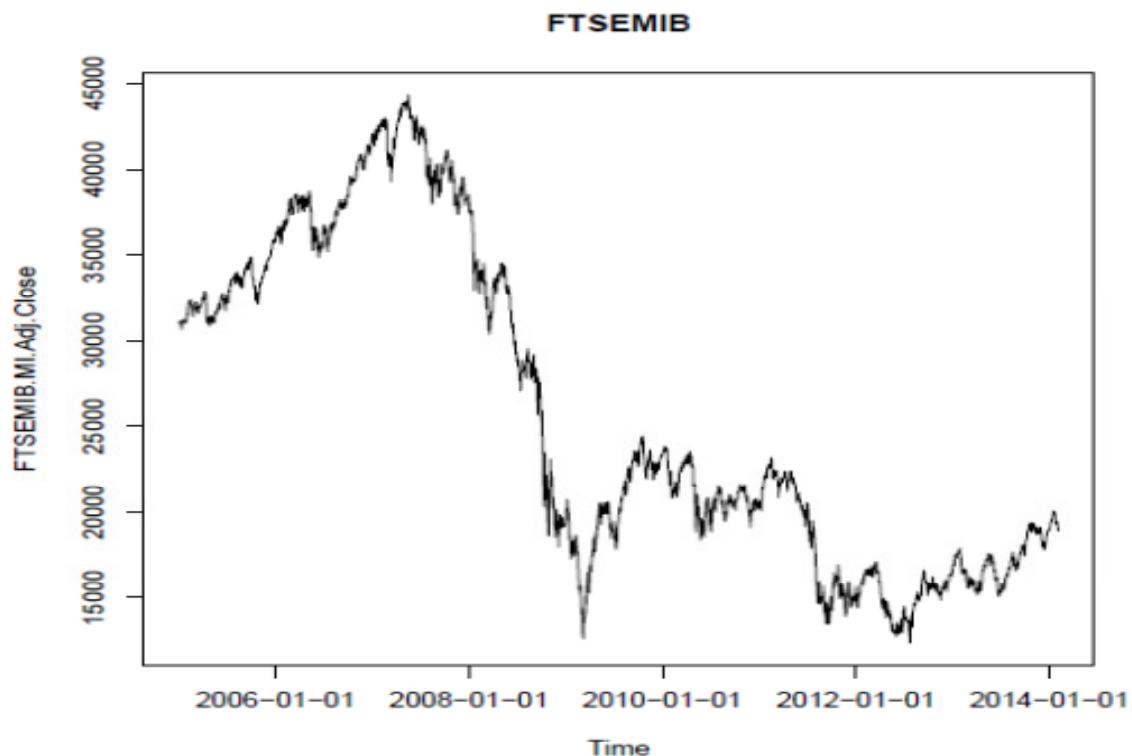
Posto infatti $\lambda = 0.4$, per un periodo di tempo dall'anno 2006 all'anno 2013, si prendano come esempio il numero di connessioni attive che 10 titoli tra quelli appartenenti all'indice FTSEMIB hanno con il resto del mercato azionario²⁸:



Si possono notare fin da subito svariate osservazioni grazie a questa analisi dinamica di applicazione della tecnica LASSO : anzitutto, i 10 titoli azionari, pur appartenendo a settori merceologici completamente diversi tra loro (bancario, assicurativo, telecomunicazioni, costruzione, tecnologia e servizi) presentano un andamento molto simile tra loro.

²⁸ il numero massimo di connessioni, in ogni tempo t , è di 27 poiché per ragioni di arco temporale di misurazione e di facilità di rilevamento, è stato possibile effettuare l'analisi su 28 titoli appartenenti al mercato FTSEMIB.

Ciò significa che in determinati periodi temporali, il numero di connessioni tra i singoli titoli e tutti gli altri titoli appartenenti al medesimo mercato, aumentano o diminuiscono simultaneamente, come se l'andamento generale del mercato azionario influenzi fortemente le connessioni tra i titoli. Ciò è dimostrato se si analizza il seguente grafico, in cui si mostra l'andamento, nello stesso arco temporale, dell'indice FTSEMIB, che rappresenta l'andamento del mercato azionario globale italiano :



Da questa analisi si nota che i periodi di crisi del mercato azionario, ovvero i due macro-periodi in cui il grafico cala vistosamente (nel 2008 e, in minor parte, nel 2011) coincidono esattamente con quei periodi in cui i singoli titoli azionari rappresentati dai grafici precedenti, intensificavano il numero di connessioni con il mercato. Quando c'è una situazione di crisi finanziaria, dunque, i rendimenti dei titoli azionari risultano essere molto più correlati tra loro, ovvero vedono aumentare il valore della covarianza tra essi.

L'applicazione della tecnica LASSO può quindi rappresentare un'interessante chiave interpretativa dell'intero mercato azionario : quando l'andamento del mercato è *stazionario* le connessioni tra i titoli sono relative al singolo business di ogni titolo (per esempio si noti come un titolo bancario come UCG mantenga sempre e comunque un numero elevato di connessioni con il resto del mercato) mentre quando il mercato azionario entra in crisi, il numero di connessioni tra i titoli aumenta.

Un'ulteriore spunto di riflessione può essere dato dalle tempistiche di reazione alla crisi globale. Preso l'anno 2007, in cui la crisi finanziaria risultava essere ancora in fase embrionale, mentre molti titoli azionari, appartenenti per lo più a settori energetici e bancari, mantenevano un numero basso di connessioni con il mercato, altri titoli come Autogrill, Buzzi Unicem e Finmeccanica vedevano già aumentare il numero delle loro interconnessioni con il mercato azionario, in vista della seguente crisi finanziaria.

Altri titoli come Eni ed Enel, ma anche i grandi gruppi bancari rappresentati da Intesa San Paolo e Unicredit, invece, alla lenta ripresa del mercato dopo la seconda crisi dell'indice FTSEMIB, ovvero nell'ultimo anno 2013, reagiscono più velocemente, interrompendo bruscamente molte connessioni con il resto del mercato azionario.

CAP. 5

SORVEGLIANZA DELLA MATRICE DI VARIANZE E COVARIANZE

5.1 Financial Surveillance

Come già visto nell'ultimo paragrafo del capitolo precedente, l'analisi effettuata nel corso di questo elaborato può essere estesa ad una visione dinamica nel tempo del mercato azionario, in cui effettivamente ad ogni tempo $t+1$ subentrano nuove osservazioni riguardanti la serie storica dei rendimenti di ogni singolo titolo azionario.

Non solo i pesi del portafoglio subiscono cambiamenti dovuti alle nuove situazioni di mercato ma anche la struttura delle dipendenze tra i titoli e, di conseguenza, la matrice Σ di varianze e covarianze che ne sintetizza l'informazione.

L'idea di questo capitolo conclusivo è quella di proporre un "ponte" tra le tecniche di stima mostrate e il concetto di Sorveglianza statistica in ambito finanziario.

L'obiettivo sarà quello di identificare nel minor tempo possibile l'istante temporale in cui la struttura della matrice di varianze e covarianze Σ subisca un brusco cambiamento²⁹, definito *punto di rottura*, il quale identifica una diversa struttura del portafoglio azionario.

Nella recente letteratura si è verificato un interesse sempre più crescente riguardo la necessità di metodi di sorveglianza, in diversi ambiti. Tutti questi settori hanno in comune una metodologia statistica. Nel *controllo statistico* si osserva una serie storica, con lo scopo di rilevare un importante cambiamento nel processo sottostante i dati il prima possibile, dopo il verificarsi di tale cambiamento, al fine di operare strategie decisionali tempestive.

Nella maggior parte dei casi, vengono dunque ampiamente utilizzati i metodi di SPC (controllo statistico di processo) per rilevare un brusco cambiamento nel caso di osservazioni sequenziali, ovvero nel caso in cui i dati vengono continuamente rilevati nel tempo e vengono analizzati ripetutamente.

²⁹ noto in letteratura anglosassone come "change point".

Questo è ciò che contraddistingue la Sorveglianza statistica : non si prende in considerazione un insieme fisso di dati, ma un numero crescente di osservazioni. Quindi non sappiamo se al tempo t un cambiamento è successo e quando sia successo, ma ripetendo la procedura iterativamente si cerca di identificare il prima possibile un *change point* del processo in una situazione in cui nuovi dati vengono osservati ad ogni istante temporale.

L'applicazione dei metodi di Sorveglianza statistica in ambito finanziario, prende il nome di *Financial Surveillance*, che ha recentemente acquisito popolarità nella letteratura moderna, grazie soprattutto agli studi di Frisén (2008), la quale pone l'attenzione sul monitoraggio, attraverso carte di controllo multivariate, dei pesi ω_i di un portafoglio azionario.

Una proposta alternativa è quella di effettuare una Sorveglianza direttamente alla matrice di varianze e covarianze del portafoglio, alla luce della nuova metodologia di stima penalizzata analizzata negli scorsi capitoli, attraverso la costruzione di opportune carte di controllo multivariate.

Sia Y una variabile casuale p -variata $(y_1, y_2, \dots, y_p) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ dove p è il numero di titoli azionari presenti in portafoglio. La costruzione di carte di controllo multivariate può essere effettuata per monitorare la media del processo, ovvero se il vettore delle medie nel tempo si discosta da un prefissato valore $\boldsymbol{\mu}_0$, oppure la variabilità del processo, se la matrice di varianze e covarianze si discosta da una struttura prefissata $\boldsymbol{\Sigma}_0$, oppure entrambe le caratteristiche analizzate congiuntamente, costruendo per esempio carte di controllo basate sulla statistica T^2 di Hotellig :

$$T^2 = n (\bar{Y} - \boldsymbol{\mu}_0) \mathbf{S}^{-1} (\bar{Y} - \boldsymbol{\mu}_0)'$$

dove $\bar{Y}$ è la media campionaria, $\mathbf{S}$ è la matrice di varianze e covarianze campionaria ed n è la numerosità del campione.

I metodi più noti in letteratura per effettuare controllo di processo multivariato, sono basati sull'estensione delle procedure di controllo univariate, attraverso la costruzione di carte di controllo Shewhart, Cusum ed Ewma, adattate al caso multivariato.

Le carte di controllo multivariate basate sulla statistica T^2 di Hotelling rappresentano un'estensione delle carte Shewhart univariate, poiché utilizzano le informazioni derivanti dal solo campione corrente, risultando così poco sensibili nel caso di piccole variazioni nel tempo, mancando di una cosiddetta *memoria* del processo³⁰.

Crosier (1988) ha proposto diverse procedure CUSUM di tipo multivariato (MCUSUM), mentre Lowry et al. (1992) furono tra i primi ad estendere al caso multivariato (MEWMA) le carte di controllo di tipo EWMA.

In tutti i casi, le procedure di controllo multivariato prevedono due fasi : la prima fase ha come obiettivo quello di stabilire i limiti di controllo, ovvero stabilire quali siano i valori per i quali il processo (o la serie storica) è da considerarsi “in controllo”, attraverso lo studio di alcuni campioni cosiddetti *preliminari*. Nella seconda fase, si effettua il controllo vero e proprio, che coincide con la sorveglianza di una statistica test calcolata ad ogni tempo t attraverso il *nuovo* campione multivariato rappresentante i rendimenti dei titoli azionari.

In ogni caso, si ricorda che qualsiasi procedura di controllo multivariato si intenda utilizzare, un approccio molto utile per l'interpretazione di un segnale di “fuori controllo” è quello di affiancare alla carta di controllo p -variata, p carte di controllo univariate per stabilire quali delle p variabili abbiano provocato tale situazione. L'approccio univariato può tuttavia risultare fuorviante, motivo per cui in letteratura sono state suggerite diverse procedure come la scomposizione della T^2 di Hotelling o l'analisi delle componenti principali³¹, che sono combinazioni lineari delle variabili originarie, ma che risultano essere incorrelate tra loro.

La proposta del presente capitolo è quella di fornire una breve illustrazione di alcune metodologie di controllo multivariato presenti in letteratura adattandole al contesto finanziario, in cui la matrice di varianze e covarianze non sia però quella campionaria, ma quella stimata attraverso l'algoritmo LASSO.

³⁰ cfr. Montgomery (2005)

³¹ Jackson (1980)

5.2 Sorveglianza della matrice di varianze e covarianze dei rendimenti dei titoli di un portafoglio azionario tramite la tecnica LASSO

Nonostante la recente letteratura in ambito finanziario abbia spesso focalizzato l'attenzione sulle carte di controllo multivariate per il monitoraggio dei pesi del portafoglio ottimo, pochi autori hanno sviluppato tecniche di monitoraggio della matrice di varianze e covarianze di un portafoglio azionario.

Un primo approccio volto a controllare nel tempo una matrice di varianze e covarianze fu proposto da Montgomery e Wadsworth nel 1972, la cui idea fu quella di sintetizzare l'intera matrice attraverso la stima del determinante. Le prime carte di controllo furono quindi tracciate studiando la distribuzione dei primi due momenti centrali di $|S|$, supposto che gran parte della sua distribuzione di probabilità sia contenuta nell'intervallo $E[|S|] \pm 3\sqrt{\text{Var}[S]}$.

Nel corso degli anni, l'utilizzo di carte di controllo multivariate per la matrice Σ , ebbe un notevole sviluppo metodologico nel 1985 grazie alla proposta di F.B. Alt, in cui la distribuzione sottostante i dati campionari si assume essere una Normale p-variata con vettore delle medie μ e matrice di varianze e covarianze Σ .

La proposta di Alt riguarda la seconda fase del controllo. Nella prima fase si assume che la variabilità del processo multivariato sia pari a un determinato Σ_0 . Successivamente alla prima fase, si costruisce una statistica test K_i basata sul rapporto di verosimiglianza³², per ogni tempo i , che rappresenta l'estensione al caso multivariato delle carte di controllo univariate per la varianza ed è data da :

$$K_i = -pn + pn \ln(n) - n \ln\left(\frac{|A_i|}{|\Sigma_0|}\right) + \text{tr}(\Sigma_0^{-1}A_i)$$

dove $A_i = (n-1)S_i$ ed S_i è la matrice di varianze e covarianze campionaria per ogni tempo i .

L'espressione precedente, in termini di S_i , si dimostra essere equivalente³³ a :

$$K_i = -(n-1) \left[p + \ln\left(\frac{|S_i|}{|\Sigma_0|}\right) - \text{tr}(\Sigma_0^{-1}S_i) \right]$$

³² derivante dai risultati contenuti in Anderson (1958).

³³ cfr. Li, Wang e Yeh (2012)

di cui ci si serve per effettuare, ad ogni istante i , un test d'ipotesi del tipo :

$$\begin{cases} H_0 : \Sigma = \Sigma_0 \\ H_1 : \Sigma \neq \Sigma_0 \end{cases}$$

nel quale si accetta l'ipotesi nulla, ovvero che il processo è "in controllo", se :

$$K_i < \chi^2_{\frac{p(p+1)}{2}, \alpha}$$

Si noti che, quando $\mathbf{S}_i = \Sigma_0$ l'espressione di K_i diventa:

$$K_i = -(n-1)[p + 0 - p] = 0$$

A partire dalla proposta di Alt, la letteratura più recente ha sviluppato nuove metodologie di controllo multivariato riguardanti la matrice di varianze e covarianze, anche nel caso essa venga penalizzata tramite un algoritmo di tipo LASSO.

Tra le altre, Li, Wang e Yeh (2012) pongono come obiettivo quello di testare l'ipotesi per cui :

$$\begin{cases} H_0 : \Sigma = \mathbf{I}_p \\ H_1 : \Sigma \neq \mathbf{I}_p \end{cases}$$

dove $\mathbf{I}_p$ è la matrice identità p -dimensionale. Quindi si presuppone che la struttura della matrice di varianze e covarianze sia "in controllo" quando presenta valore unitario sulla diagonale principale e zero sugli elementi extra-diagonali, ovvero nel caso di variabili standardizzate.

In tal caso, il rapporto di log-verosimiglianza può essere riscritto come :

$$\Lambda = tr(\mathbf{S}) - tr(\Theta \mathbf{S}) + \ln |\Theta|$$

che coincide con la statistica K_i presentata da Alt se la matrice $\Theta = \Sigma^{-1}$ viene stimata con la massima verosimiglianza campionaria : $\Theta = \frac{n}{n-1} \mathbf{S}^{-1}$.

Nel caso in cui, invece, la stima di Σ , e quindi della sua inversa Θ , viene effettuata risolvendo un algoritmo di penalizzazione come quello proposto dalla tecnica LASSO con parametro λ , si ottiene una stima della statistica :

$$\Lambda_\lambda = tr(\mathbf{S}) - tr(\mathbf{\Theta}_\lambda \mathbf{S}) + \ln |\mathbf{\Theta}_\lambda|$$

che rappresenta un “fuori controllo” quando : $\Lambda_\lambda > h_\lambda$ dove h_λ viene ottenuto tramite simulazioni, per assicurare un determinato livello di ARL³⁴, ovvero per ottenere un segnale di fuori controllo il prima possibile, quando esso accade.

Nel controllo statistico multivariato, e quindi anche nel contesto finanziario oggetto di questo lavoro, l’analisi avviene però in due fasi e nella prima fase è necessario individuare una struttura di $\mathbf{\Sigma}_0$ alla quale fare riferimento, per un test d’ipotesi del tipo :

$$\begin{cases} H_0 : \mathbf{\Sigma} = \mathbf{\Sigma}_0 \\ H_1 : \mathbf{\Sigma} \neq \mathbf{\Sigma}_0 \end{cases}$$

L’obiettivo infatti, nel nostro caso, non è quello di confrontare la matrice di varianze e covarianze stimata tramite la tecnica LASSO con la matrice identità, ma con una struttura prefissata che indichi quando il legame tra i titoli presenti in portafoglio è da considerarsi “in controllo”.

La nostra proposta è quella di assumere come $\mathbf{\Sigma}_0$ di riferimento la matrice di varianze e covarianze stimata tramite il modello a un fattore, cui fa riferimento il paragrafo 1.10, in cui si *depura* l’effetto del mercato a cui tutti i titoli appartengono, ovvero l’effetto determinato dall’andamento dell’indice FTSEMIB. Questa metodologia permette di confrontare la nuova stima penalizzata della matrice di varianze e covarianze $\mathbf{\Sigma}_\lambda$ con una struttura $\mathbf{\Sigma}_0$ nella quale gli elementi diagonali rappresentano le misure di rischio intrinseche di ogni singolo titolo, mentre gli elementi extradiagonali sono nulli solo per quei titoli che risultano essere effettivamente incorrelati nella realtà, come si è visto ad esempio nel paragrafo 1.10 riguardo il titolo FIAT rispetto ad altri titoli appartenenti al settore bancario.

Il confronto tra la nuova matrice di varianze e covarianze penalizzata e quella stimata attraverso la depurazione dall’effetto del mercato è ciò che avviene anche nella maggior parte delle tecniche di stima *shrinkage*, in cui per ottenere una stima della matrice di varianze e covarianze penalizzata si effettua una combinazione lineare tra quella ottenuta tramite il campione e quella derivante da una stima SIM (*Single Index Model*) in modo tale da avvicinarsi al modello depurato dall’effetto del FTSEMIB.

³⁴ Average Run Length, numero atteso di campioni che devono essere osservati prima che lo scostamento dal valore di riferimento venga individuato.

Una volta determinata la matrice Σ_0 di riferimento, per passare da un controllo del tipo :

$$\begin{cases} H_0 : \Sigma = \mathbf{I}_p \\ H_1 : \Sigma \neq \mathbf{I}_p \end{cases} \quad (1)$$

al generico controllo derivante dal test d'ipotesi :

$$\begin{cases} H_0 : \Sigma = \Sigma_0 \\ H_1 : \Sigma \neq \Sigma_0 \end{cases} \quad (2)$$

è necessario pre-moltiplicare e post-moltiplicare la matrice $\Sigma_0^{-\frac{1}{2}}$ alla matrice Σ in maniera tale da utilizzare la metodologia di controllo proposta da Li, Wang e Yeh (2012) che invece prevedeva un controllo multivariato rispetto alla matrice identità $\mathbf{I}_p$.

Quindi, in sintesi, la procedura per la costruzione di una carta di controllo multivariata per la matrice di varianze e covarianze Σ prevede :

1. identificazione di una matrice Σ_0 “in controllo” (fase 1).
2. stima della matrice Σ_λ penalizzata, seguendo l’algoritmo LASSO.
3. controllo della matrice Σ_λ rispetto alla matrice Σ_0 (fase 2) che prevede :

- a. trasformazione in : $\Sigma = \left(\Sigma_0^{-\frac{1}{2}} \right) \Sigma_\lambda \left(\Sigma_0^{-\frac{1}{2}} \right)'$

- b. calcolo della statistica : $\Lambda = tr(\mathbf{S}) - tr(\Theta \mathbf{S}) + \ln |\Theta|$ dove : $\Theta = \Sigma^{-1}$

- c. stima del limite di controllo h_λ .

4. costruzione della carta di controllo e identificazione dei punti di “fuori controllo”.

Un secondo interessante approccio, alternativo al precedente, è quello proposto da Maboudou-Tchao, Diawara (2012) in cui si ottiene una carta di controllo multivariata per la matrice Σ a partire da una procedura di tipo MEWMC (*Multivariate Exponentially Weighted Moving Covariance Matrix*). Posto che, sotto l'ipotesi di essere “in controllo”, ogni osservazione multivariata $\mathbf{x}_i$ sia derivante da una distribuzione $N(\boldsymbol{\mu}, \Sigma_0)$, si effettua anche

in questo caso una standardizzazione tale per cui da $\mathbf{x}_i$ si passa a $\mathbf{u}_i = \Sigma_0^{-\frac{1}{2}} (\mathbf{x}_i - \boldsymbol{\mu})$ in modo tale che, quando il processo è “in controllo”, $\mathbf{u}_i$ sia $N(\mathbf{0}, \mathbf{I}_p)$.

La carta di controllo MEWMC si ottiene ricorsivamente, per ogni tempo i , a partire dalla matrice identità $\mathbf{K}_0 = \mathbf{I}_p$ e successivamente :

$$\mathbf{K}_i = (1 - \lambda)\mathbf{K}_{i-1} + \lambda \mathbf{u}_i \mathbf{u}_i'$$

dove $0 < \lambda < 1$ è definito parametro di smorzamento.

Nel caso in cui, al tempo i , la stima della matrice di varianze e covarianze $\mathbf{\Sigma}$, ottenuta tramite la tecnica LASSO, si discosti dal valore “in controllo” dato da $\mathbf{\Sigma}_0$, la trasformazione data da $\mathbf{u}_i = \mathbf{\Sigma}_0^{-\frac{1}{2}} (\mathbf{x}_i - \boldsymbol{\mu})$ si allontanerebbe dalla situazione standard di $\mathbf{u}_i \sim N(\mathbf{0}, \mathbf{I}_p)$.

La statistica-test calcolata per costruire la carta di controllo è data quindi da :

$$C_i = tr(\mathbf{K}_i) - Ln|\mathbf{K}_i| - p$$

la quale identifica la regione di fuori controllo se $C_i > h$ dove h viene ottenuto tramite simulazioni per assicurare un determinato livello di ARL.

Se invece $\mathbf{K}_i = \mathbf{I}_p$ vale che :

$$C_i = p - 0 - p = 0$$

Gli autori³⁵ propongono infine una carta di controllo in cui l’algoritmo LASSO venga applicato direttamente alla stima della matrice $\mathbf{U}_i = \mathbf{u}_i \mathbf{u}_i'$ ottenendo dunque una procedura denominata LEWMC (*Lasso Multivariate Exponentially Weighted Moving Covariance Matrix*). In tale procedura è sufficiente standardizzare ogni osservazione $\mathbf{x}_i$ in $\mathbf{u}_i = \mathbf{A}(\mathbf{x}_i - \boldsymbol{\mu})$ dove $\mathbf{A}$ è una matrice tale per cui $\mathbf{A} \mathbf{\Sigma} \mathbf{A}' = \mathbf{I}_p$ in modo tale da stimare la matrice $\mathbf{V}_i$ (penalizzando $\mathbf{U}_i = \mathbf{u}_i \mathbf{u}_i'$) e confrontare l’andamento di $\mathbf{K}_i = (1 - \lambda)\mathbf{K}_{i-1} + \lambda \mathbf{V}_i$ con la matrice identità, attraverso la statistica $C_i = tr(\mathbf{K}_i) - Ln|\mathbf{K}_i| - p$.

³⁵ Maboudou-Tchao, Diawara (2012)

5.3 Costruzione di una carta di controllo : un'applicazione

A partire dalla proposta di Maboudou-Tchao e Diawara (2012) tramite il procedimento riportato in Appendice, è possibile costruire una carta di controllo della matrice Σ di varianze e covarianze di un portafoglio azionario per individuare nel minor tempo possibile l'istante a partire dal quale si può assumere che si sia modificata la struttura della matrice.

Si consideri un portafoglio composto dai quattro titoli azionari riportati nell'Esempio 2 : Intesa San Paolo, Unicredit, Fiat e Mediaset e si osservino i rendimenti dei titoli a partire da Gennaio 2012 fino a Marzo 2014.

Si prenda in esame il test d'ipotesi :

$$\begin{cases} H_0 : \Sigma = \Sigma_0 \\ H_1 : \Sigma \neq \Sigma_0 \end{cases}$$

Il controllo avviene in due fasi.

Nella prima fase, composta a titolo esemplificativo dai primi 100 giorni presi in considerazione (*training set*), si identifica la struttura di riferimento Σ_0 stimata tramite il modello a un fattore, in cui si *depura* l'effetto del mercato a cui i titoli appartengono, e si individua il limite superiore di controllo, ovvero la soglia oltre alla quale il processo è da considerarsi "fuori controllo", in quanto la struttura di Σ si discosta significativamente³⁶ dalla matrice Σ_0 di riferimento.

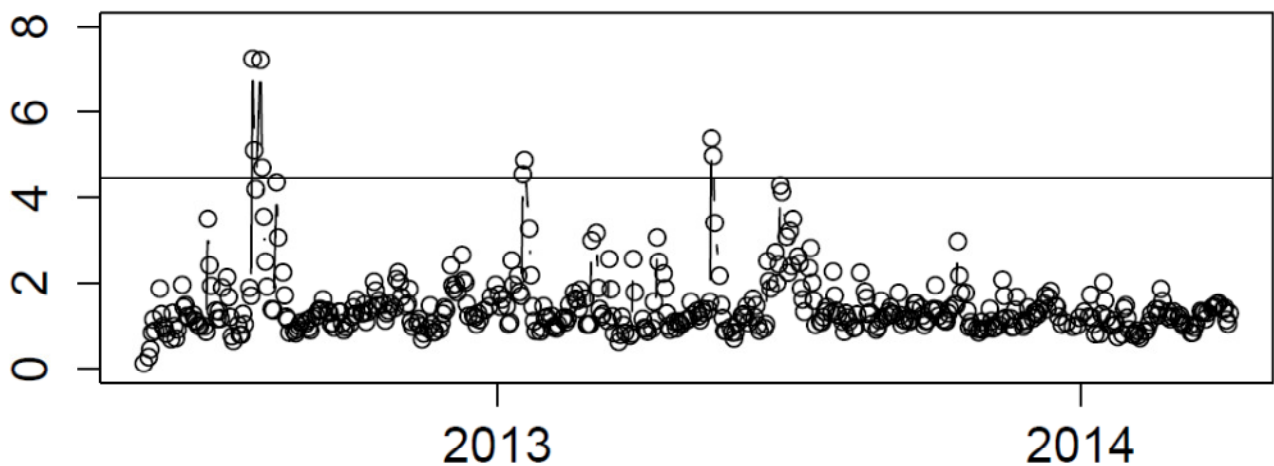
Nella seconda fase avviene il controllo vero e proprio secondo la seguente procedura, per ogni istante temporale i :

1. standardizzazione delle osservazioni : il vettore x_i p -dimensionale, dove p è il numero di titoli componenti il portafoglio, viene trasformato in $u_i = \Sigma_0^{-\frac{1}{2}} (x_i - \mu_0)$ con Σ_0 e μ_0 stimati nella fase 1.
2. costruzione della matrice $U_i = u_i u_i'$
3. stima della matrice V_i applicando l'algoritmo di penalizzazione LASSO alla matrice U_i
4. stima della matrice $K_i = (1 - \lambda)K_{i-1} + \lambda V_i$ a partire da $K_0 = I_N$
5. costruzione della statistica-test : $C_i = tr(K_i) - Ln|K_i| - p$

³⁶ nell'esempio riportato in Appendice è stata presa in considerazione una probabilità $(1-\alpha) = 95\%$

Oltre alla scelta del numero di osservazioni componenti il *training set* e alla scelta del limite di controllo effettuati nella prima fase, le quantità incognite coinvolte nella procedura sono la penalizzazione dell'algoritmo LASSO, che può essere individuato come presentato nel paragrafo 4.2, e il parametro λ che identifica il peso da allocare alle nuove osservazioni nella stima della matrice $\mathbf{K}_t$.

Fissato un livello di penalizzazione di 0.2 e un parametro $\lambda = 0.3$ è dunque possibile costruire una carta di controllo come riportato nella seguente figura :



In questo esempio, per ogni istante temporale successivo al *training set*, si verifica che la statistica-test C_t supera il limite di controllo (fissato a 4.45) in quegli istanti temporali in cui la matrice di varianze e covarianze $\mathbf{\Sigma}$ si discosta significativamente dalla matrice di riferimento $\mathbf{\Sigma}_0$, individuata nel *training set*.

CONCLUSIONI E SVILUPPI FUTURI

L'obiettivo principale del presente lavoro è stato quello di proporre in ambito finanziario una nuova metodologia di stima della matrice di varianze e covarianze dei rendimenti di un portafoglio azionario tramite una tecnica nota in letteratura come LASSO, che permetta, oltre ad effettuare un'operazione di penalizzazione (*shrinkage* in inglese), anche una *variable selection* tra i titoli.

L'innovazione proposta è stata quindi quella di effettuare un'operazione di *portfolio selection*, grazie alla quale è stato possibile migliorare la comprensione della composizione del portafoglio in un'ottica di allocazione ottimale delle risorse.

Inoltre è stato utilizzato uno strumento legato alla teoria dei grafi che proprio tramite un'operazione di *variable selection* ha potuto mostrare le interdipendenze tra i titoli presenti in portafoglio, sia in un'ottica statica che in un'ottica continua nel tempo, al fine di individuare le dinamiche sottostanti il mercato azionario, tramite il numero di connessioni che ogni singolo titolo ha con il resto del mercato.

La rilevanza e l'attualità dell'argomento in questione è confermata dalla crescente attenzione rivolta alle tecniche di stima penalizzata, che molti autori hanno introdotto negli ultimi anni in contrapposizione alle più note tecniche di stima campionaria e di massima verosimiglianza.

La parte conclusiva dell'elaborato mostra infine un collegamento tra queste tecniche e l'argomento di *Financial Surveillance* trattato in particolare negli studi di M.Frisen. L'obiettivo in questo caso è stato quello di effettuare sorveglianza statistica applicata direttamente alla matrice penalizzata di varianze e covarianze dei rendimenti, tramite carte di controllo multivariate che permettano di individuare nel minor tempo possibile un punto di rottura (*change point*), ovvero il momento in cui la struttura di dipendenza sottostante i titoli subisca un forte cambiamento, che porta ad una diversa composizione del portafoglio azionario.

Un interessante sviluppo futuro degli argomenti trattati può prendere spunto dal fatto che la letteratura moderna in ambito finanziario si stia sempre più orientando verso un'allocazione ottimale di portafoglio a quattro dimensioni, ovvero che tenga in considerazione anche lo studio dell'indice di asimmetria e di curtosi dei singoli titoli azionari.

Inoltre, nonostante in questo lavoro sia stata spesso ipotizzata l'assunzione di normalità dei rendimenti azionari, tale ipotesi può non essere sempre verificata nella realtà e si rende dunque necessario un allontanamento dall'assunto di normalità dei rendimenti. Una possibilità è quella di servirsi delle cosiddette Copula.

Una Copula è una funzione C che permette di costruire una distribuzione multivariata partendo da due o più variabili marginali, ipotizzando una struttura di dipendenza tra di esse. Nel caso bivariato, date le funzioni di ripartizioni marginali $F(x)$ ed $F(y)$, la funzione di ripartizione congiunta è quindi data da :

$$H(x, y) = C [F(x), F(y)]$$

La Copula definisce quindi il tipo di dipendenza tra le variabili marginali.

Per esempio, ipotizzando una dipendenza di tipo *lineare* tra le distribuzioni marginali, e utilizzando una Copula gaussiana del tipo :

$$C(u, v) = \int_{-\infty}^{\phi^{-1}(u)} \int_{-\infty}^{\phi^{-1}(v)} \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp \left\{ -\frac{s_1^2 - 2\rho s_1 s_2 + s_2^2}{2(1-\rho^2)} \right\} ds_1 ds_2$$

è possibile costruire una distribuzione multivariata con una struttura di dipendenza di tipo normale, indipendentemente dalla distribuzione delle singole marginali. In questo modo è possibile non soddisfare l'ipotesi di normalità dei log-rendimenti dei titoli azionari.

Costruendo una Copula gaussiana con funzioni di ripartizioni marginali empiriche qualsiasi, si ottiene quindi una stima semi-parametrica della distribuzione multivariata, con dipendenza di tipo lineare. A questo punto la stima di massima verosimiglianza della matrice di varianze e covarianze Σ ottenuta tramite la Copula gaussiana può essere impiegata nell'algoritmo LASSO che è stato presentato in questo lavoro.

Tra le *Copule ellittiche*, infine, ovvero quella famiglia di Copule caratterizzate da un importante legame con il coefficiente di correlazione lineare ρ , grazie alle quali è possibile stimare una matrice di varianze e covarianze Σ , si noti che l'utilizzo di una Copula T permetterebbe anche di valutare l'eventuale dipendenza tra le code, che la Copula gaussiana suppone indipendenti.

APPENDICE

Script di R

Nella presente appendice conclusiva vengono riportati i listati del software R utilizzati nelle applicazioni mostrate nel corso dell'elaborato.

PARAGRAFO 2.3 ³⁷

```
library(quadprog)
library(glasso)
library(fImport)
#### codici su yahoo finance dei titoli in ftse.mib
Cod_FTSE_MIB <- c("ISP.MI","UCG.MI","F.MI","MS.MI")
#### serie ftsemib
FTSEMIB <- yahooSeries("FTSEMIB.MI")
portaf <- FTSEMIB[,6]
for(i in 1:length(Cod_FTSE_MIB))
{
  titolo <- yahooSeries(Cod_FTSE_MIB[i])
  portaf <- (merge(portaf,titolo[,6]))
}
portaf <- portaf[,-1]
portaf_ret <- returns(portaf)
names(portaf_ret) <- Cod_FTSE_MIB
R <- portaf_ret

mufree = 0.0001
matrout <- NULL
for(j in 1:2)
{
  mean_vect = apply(R,2,mean)
  cov_mat = cov(R)
  if(j == 2) {
    a<-glasso(cov_mat, rho=.0001)
    cov_mat <- a$w }
  solve(cov_mat)
  sd_vect = sqrt(diag(cov_mat))
  Amat = cbind(rep(1,4),mean_vect,diag(1,nrow=4))
  muP = seq(min(mean_vect)+.0001,max(mean_vect)-.0001,length=300)
  sdP = muP
  weights = matrix(0,nrow=300,ncol=4)
  for (i in 1:length(muP)) {
    bvec = c(1,muP[i],rep(0,4))
    result = solve.QP(Dmat=2*cov_mat,dvec=rep(0,4),Amat=Amat,bvec=bvec,meq=2)
    sdP[i] = sqrt(result$value)
    weights[i,] = result$solution }
```

³⁷ si veda anche Ruppert (2011)


```

sharpe =( muP-mufree)/sdP
ind = (sharpe == max(sharpe))
options(digits=3)
ind2 = (sdP == min(sdP))
ind3 = (muP > muP[ind2])
options(scipen=0, digits=5 )
if(j==1){
  plot(sdP,muP, xlim=c(0,max(sdP*1.1)), ylim=c(min(muP*0.8),max(muP*1.5)))
  lines(c(0,2),mufree+c(0,2)*(muP[ind]-mufree)/sdP[ind],lwd=2,lty=2)
  points(sdP[ind],muP[ind],cex=2,pch="*")
  ind2 = (sdP == min(sdP))
  points(sdP[ind2],muP[ind2],cex=2,pch="+", col="red")
  muini_front=muP[ind]
  matrout <- rbind(matrout,c(weights[ind,], sdP[ind], muP[ind]))
  muini_min=muP[ind2]
  matrout <- rbind(matrout,c(weights[ind2,], sdP[ind2], muP[ind2])) }
if(j==2){
  points(sdP,muP, xlim=c(0,max(sdP*1.1)), ylim=c(min(muP*0.8),max(muP*1.5)), col="green")
  points(sdP[ind],muP[ind],cex=2,pch="*")
  ind2 = (sdP == min(sdP))
  points(sdP[ind2],muP[ind2],cex=2,pch="+", col="blue")
  lines(c(0,2),mufree+c(0,2)*(muP[ind]-mufree)/sdP[ind],lwd=1,lty=1, col="yellow")
  matrout <- rbind(matrout,c(weights[ind,], sdP[ind], muP[ind]))
  ind= which(muP==muini_front)
  matrout <- rbind(matrout,c(weights[ind2,], sdP[ind2], muP[ind2])) }
}
matrout <- data.frame(matrout)
names(matrout) <- c("w1","w2","w3","w4","sdP","muP")
rownames(matrout) <- c("opt front tang","opt min var","opt front tang LASSO","opt min var LASSO")
print(matrout)

```

PARAGRAFO 3.2³⁸

```

library(glasso)
library(igraph)
library(graph)
library(Rgraphviz)
library(mvtnorm)
library(flmpport)
#### codici su yahoo finance dei titoli in ftse.mib
Cod_FTSE_MIB <- c("ISP.MI","UCG.MI","F.MI","MS.MI")
#### serie ftsemib
FTSEMIB <- yahooSeries("FTSEMIB.MI")
portaf <- FTSEMIB[,6]
for(i in 1:length(Cod_FTSE_MIB))
{
  titolo <- yahooSeries(Cod_FTSE_MIB[i])
  portaf <- (merge(portaf,titolo[,6]))
}

```

³⁸ si veda anche Edwards (2000)

```

portaf <- portaf[,-1]
portaf_ret <- returns(portaf)
names(portaf_ret) <- Cod_FTSE_MIB
R <- portaf_ret

cov_mat = cov(R)
diagprinc <- diag(dim(cov_mat)[1])
par(mfrow=c(3,3))
firstpass=TRUE
for(rho_step in seq(.0,.00045, length=9))
{
  a<-glasso(cov_mat, rho=rho_step)
  cov_mat_glasso <- a$w
  cov_mat_inversa_glasso <- a$wi
  rownames(cov_mat_inversa_glasso)<-colnames(cov_mat_inversa_glasso)<-colnames(cov_mat)
  print(cov_mat_inversa_glasso)
  c<-ifelse(cov_mat_inversa_glasso[,,]!=0,1,0)
  c <- c-diagprinc
  print(c)
  aa <- graph.adjacency(c)
  if(firstpass) {
    lay_ini<- layout.spring(aa)
    aa$layout <- lay_ini
    firstpass=F}
  if(!firstpass) aa$layout=lay_ini
  E(aa)$arrow.mode <- c(0,0)
  V(aa)$size <- 10
  plot(aa, main=paste("λ=" , rho_step))
}

```

PARAGRAFO 5.3³⁹

```

LEWMC.Chart <- function(x=portaf_ret,rho=.2,cl=1,lam=.3)
{
  ## x: data
  ## rho: LASSO penalty term
  ## cl: Control limit
  ## lam: Smoothing constant term
  if (rho<=0)
    stop("The tuning parameter must be a positive number.")
  if (lam<=0 || lam>=1)
    stop("The smoothing constant must be a number between 0 and 1.")
  if (cl <= 0)
    stop("The control limit must be a positive number.")
  n<-dim(x)[1]
  p<- dim(x) [2]
  sn<-diag(p)
  cn<-rep( 0 ,n)
  sample<-rep(0 ,n)
  k<-0
  cn<-rep(0,n)

```

³⁹ si veda anche Maboudou-Tchao, Diawara (2012)

```

for(i in 1:n){
  k<-k+1
  u<-x[i,]
  sample[i]<-k
  U<-as.numeric(u) %*% t(as.numeric(u))
  V<-glasso(U,rho)$w
  sn<-(1-lam)*sn +lam*V
  trsn<-sum(diag(sn))
  cn[i]<-trsn-log(det(sn))-p
}
cn
}

library(flimport)
library(glasso)
Cod_FTSE_MIB <-c("ISP.MI","UCG.MI","F.MI","MS.MI")
FTSEMIB <- yahooSeries("FTSEMIB.MI",from="2012-01-01")
portaf <- FTSEMIB[,6]
for(i in 1:length(Cod_FTSE_MIB))
{
  titolo <- yahooSeries(Cod_FTSE_MIB[i],from="2012-01-01")
  portaf <- (merge(portaf,titolo[,6]))
}
portaf_ret <- (returns(portaf)*100)
names(portaf) <- c("FTSEMIB.MI",Cod_FTSE_MIB)
names(portaf_ret) <- c("FTSEMIB.MI",Cod_FTSE_MIB)
portaf_ret <- portaf_ret[-508,]
portaf_ret <- portaf_ret[-502,]
matr_res<- NULL
for(i in 2: (length(Cod_FTSE_MIB)+1)){
  res<- lm(portaf_ret[,i]~portaf_ret[,1],data= portaf_ret)$residuals
  cbind(res,res)
  matr_res <- cbind(matr_res,res)
}
names(matr_res) <- c(Cod_FTSE_MIB)
colnames(matr_res) <- Cod_FTSE_MIB
portaf_ret <- data.frame(portaf_ret[, -1])
matr_res <- data.frame(matr_res)
names(portaf_ret)
training_set <- 100
vect_sigma_0 <- (diag(var(matr_res[1:training_set,])))^0.5
mu_0 <- apply(matr_res[1:training_set,], MARGIN=2, mean)
matr_medie <- rep(1,training_set) %*% t(mu_0)
centrata_0 <- matr_res[1:training_set,]-matr_medie
matr_sqm <- rep(1,training_set) %*% t(1/vect_sigma_0)
matr_sotto_H0 <- (centrata_0 * matr_sqm)
cn_H0 <- LEWMC.Chart(x=matr_sotto_H0,rho=.2,cl=1,lam=.3)
cl_H0 <- quantile(cn_H0, probs=.95)
matr_medie <- rep(1,dim(matr_res)[1]-training_set) %*% t(mu_0)
centrata_1 <- matr_res[(training_set+1):dim(matr_res)[1,]-matr_medie
matr_sqm <- rep(1,dim(matr_res)[1]-training_set) %*% t(1/vect_sigma_0)
matr_sotto_H1 <- (centrata_1 * matr_sqm)
cn_H1 <- LEWMC.Chart(x=matr_sotto_H1,rho=.2,cl=cl_H0,lam=.3)
cn_H1 <- (ts(cn_H1))
plot(as.Date(rownames(matr_res)[(trainig_set+1):dim(matr_res)[1]]),cn_H1,type="b", ylim=c(0,8))
abline(h=cl_H0)

```

BIBLIOGRAFIA

1. Alt (1985) **Multivariate Quality Control**. In Encyclopedia of Statistical Sciences, vol. 6, p. 110-122. John Wiley, New York.
2. Banerjee et al. (2008) **Model selection through sparse maximum likelihood estimation for multivariate gaussian or binary data**. Journal of Machine Learning Research 9, p. 485-516.
3. Benninga (2011) **Modelli finanziari, la finanza con Excel**. McGraw-Hill.
4. Black et al. (1972) **The valuation of option contracts and a test of market efficiency**. Journal of Finance, vol. 27, p. 399–418.
5. Chan, Zhang (2001) **Cumulative Sum Control Charts for Covariance Matrix**. Statistica Sinica, vol. 11.
6. Crosier (1988) **Multivariate generalizations of cumulative sum quality control schemes**. Technometrics, vol. 30.
7. Daniel, Kass (2001) **Shrinkage Estimators for Covariance Matrices**. *Biometrics*, vol. 57(4) p. 1173–1184.
8. De Miguel et al. (2013) **Size matters: Optimal calibration of shrinkage estimators for portfolio selection**. Journal of Banking and Finance, vol. 37, p. 3018–3034.
9. Draper, Smith (1998) **Applied Regression Analysis, 3rd Edition**. Wiley.
10. Edwards (2000) **Introduction to graphical modelling**. Springer.
11. Elton, Gruber (1981) **Modern Portfolio theory and investment analysis**. Wiley.
12. Friedman, Hastie, Tibshirani (2008) **Sparse inverse covariance estimation with the graphical lasso**. Biostatistics, vol. 9, issue 3, p. 432-441.
13. Frisen (2007) **Financial Surveillance**. Wiley.
14. Hull (2012) **Options, futures and other derivatives**. Pearson, 8th edition, cap.13-14.
15. Jackson (1980) **Principal components and factor analysis**. Journal of Quality Technology, 12.
16. Kourtis et al. (2012) **Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix**. Journal of Banking and Finance, vol. 36, p. 2522–2531.
17. Ledoit, Wolf (2003) **Improved estimation of the covariance matrix of stock returns with an application to portfolio selection**. Journal of Empirical Finance 10, p.603-621
18. Li et al. (2012) **Monitoring the covariance matrix via penalized likelihood estimation**. IIE Transactions, vol. 45, issue 2.

19. Lowry et al. (1992) **A multivariate exponentially weighted moving average control chart.** *Technometrics*, vol. 34.
20. Maboudou-Tchao, Diawara (2012) **A lasso chart for monitoring the covariance Matrix.** *Quality technology & quantitative management*, vol.10, p. 95-114.
21. Markowitz (1952) **Portfolio selection.** *The Journal of Finance*.
22. Markowitz (1987) **Mean-Variance Analysis in Portfolio Choice and Capital Markets.** Wiley.
23. Meinshausen e Bühlmann (2006) **High dimensional graphs and variable selection with the lasso.** *Annals of statistics*, vol. 34, p. 1436–1462.
24. Merton (1972) **An analytic derivation of the efficient portfolio frontier.** *The Journal of Financial and Quantitative Analysis*, Vol. 7, No. 4, p. 1851-1872.
25. Miller (2002) **Subset Selection in Regression. Second Edition.** Chapman&Hall/CRC.
26. Montgomery (2005) **Introduction to Statistical Quality control, Fifth Edition.** Wiley.
27. Ruppert (2004) **Statistics and Finance. An introduction.** Springer.
28. Ruppert (2011) **Statistics and Data Analysis for Financial Engineering.** Springer.
29. Schafer, Strimmer (2005) **A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics.** *Statistical Applications in Genetics and Molecular Biology*, Vol. 4, Issue 1, Art. 32.
30. Sundberg (2002) **Continuum regression.** Article for 2nd ed. of *Encyclopedia of Statistical Sciences*.
31. Theiler (2012) **The incredible shrinking covariance estimator.** *Proc. SPIE*, vol.8391.
32. Tibshirani (1996) **Regression shrinkage and selection via the Lasso.** *Journal of the Royal Statistical Society*, Vol. 58, No. 1, p. 267-288.
33. Van der Wijst (2013) **Finance, a quantitative introduction.** Cambridge University Press.
34. Wessman (1998) **Some principles for surveillance adopted for multivariate processes with a common change point.** *Communications in Statistics - Theory and Methods*, Vol. 27, Issue 5.