



Università
Ca' Foscari
Venezia

Corso di Laurea Magistrale

in Economia e Finanza

(ordinamento ex D.M. 270/2004)

Tesi di Laurea

**PSO e FWA: due metaeuristiche per la
selezione di portafogli basati sull'Entropic
Value-at-Risk come misura di rischio**

Relatore

Prof. Marco Corazza

Correlatore

Prof.ssa Martina Nardon

Laureando

Luca Marchetti

Matricola 848670

Anno Accademico

2018 / 2019

*“La misura dell’intelligenza è data dalla
capacità di cambiare quando è necessario”*

ALBERT EINSTEIN

Ringraziamenti

Desidero innanzitutto ringraziare il Prof. Corazza che, in qualità di relatore, mi ha fornito con grande disponibilità tutto l'aiuto necessario per la stesura di questa tesi.

Un sentito ringraziamento va inoltre ai miei genitori e ai miei nonni che hanno sempre creduto in me e che mi hanno costantemente motivato e sostenuto durante questi cinque anni di università.

Desidero inoltre ringraziare la società Alcedo SGR S.p.A. e tutto lo staff - in particolare Sonia - per tutto quello che mi hanno insegnato durante il percorso di formazione.

Un ringraziamento va anche allo staff della società di revisione Ecovis DCA Limited, con cui ho avuto l'opportunità di confrontarmi e crescere professionalmente durante la mia esperienza formativa a Dublino.

Infine, un ringraziamento speciale va ai miei compagni di studi ed in particolar modo a Federico, Filippo e Valentina che in questi anni mi sono stati vicini sia nei momenti felici, che in quelli difficili: sono stati per me più veri amici che semplici compagni.

Indice

Introduzione	9
1 MODERN PORTFOLIO THEORY	11
1.1 Aspetti introduttivi	11
1.2 Il processo di selezione di portafoglio	15
1.2.1 <i>Misure di rischio e rendimento</i>	<i>15</i>
1.2.2 <i>Criterio di dominanza stocastica e frontiera efficiente</i>	<i>22</i>
1.2.3 <i>Dalla teoria dell'utilità attesa al portafoglio ottimo</i>	<i>34</i>
1.3 Critiche alla Modern Portfolio Theory.....	43
2 DALLE MISURE DI RISCHIO CLASSICHE ALLE MISURE DI RISCHIO COERENTI.....	49
2.1 I limiti delle misure di rischio classiche	49
2.1.1 <i>Proprietà delle misure di rischio.....</i>	<i>51</i>
2.2 Le misure di rischio premoderne	54
2.2.1 <i>Vantaggi e limiti del Value-at-Risk.....</i>	<i>55</i>
2.3 Le misure di rischio coerenti	60
2.3.1 <i>Gli assiomi di coerenza</i>	<i>60</i>
2.3.2 <i>TCE e WCE.....</i>	<i>64</i>
2.3.3 <i>Expected Shortfall</i>	<i>66</i>
2.4 Le misure di rischio entropiche	70
2.4.1 <i>La bioeconomia</i>	<i>71</i>
2.4.2 <i>L'entropia</i>	<i>74</i>
2.4.3 <i>Entropic Value-at-Risk.....</i>	<i>79</i>
3 UN MODELLO DI SELEZIONE DI PORTAFOGLIO	87
3.1 La funzione obiettivo.....	87
3.2 Applicazione della misura di rischio.....	89
3.3 Vincoli e complessità computazionale.....	90
4 PARTICLE SWARM OPTIMIZATION E FIREWORKS ALGORITHM	97
4.1 Dalle euristiche alle metaeuristiche	97
4.1.1 <i>Le metaeuristiche.....</i>	<i>101</i>
4.2 Swarm Intelligence	104

4.3 Particle Swarm Optimization	106
4.4 La struttura della PSO	110
4.4.1 <i>L'algoritmo originale</i>	110
4.4.2 <i>Le modifiche alla PSO</i>	114
4.4.2.1 <i>Peso d'inerzia</i>	114
4.4.2.2 <i>Constriction coefficients</i>	116
4.4.2.3 <i>Fully Informed Particle Swarms (FIPS)</i>	117
4.4.3 <i>Topologia della popolazione</i>	119
4.5 Fireworks Algorithm	122
4.6 La struttura del FWA	124
4.6.1 <i>Explosion strenght</i>	125
4.6.2 <i>Explosion amplitude</i>	127
4.6.3 <i>Displacement operation</i>	129
4.6.4 <i>Mutation operator</i>	131
4.6.5 <i>Mapping rule</i>	132
4.6.6 <i>Selection strategy</i>	132
4.6.7 <i>Pseudo-codice FWA</i>	134
4.7 Caratteristiche del Fireworks Algorithm	136
4.8 Confronto tra PSO e FWA	137
5 APPLICAZIONE	141
5.1 Riformulazione del problema di selezione di portafoglio	141
5.1.1 <i>Metodo delle penalità esatte</i>	142
5.2 Due casi a confronto	145
5.3 Impostazione dei parametri del modello	151
5.3.1 <i>PSO - impostazione dei parametri</i>	152
5.3.2 <i>FWA - impostazione dei parametri</i>	154
5.4 Applicazione	154
5.4.1 <i>Confronto tra PSO e FWA</i>	155
5.4.2 <i>Analisi degli scenari</i>	156
5.4.2.1 <i>Scenari FTSE MIB</i>	157
5.4.2.2 <i>Scenari OMX Stockholm 30</i>	163
Conclusioni	170
Appendice A	172

Appendice B.....	176
Appendice C	184
Appendice D.....	187
Bibliografia	190
Sitografia	193

Introduzione

L'obiettivo di questo lavoro è quello di costruire un reale modello di selezione di portafoglio che possa fungere da concreto strumento informativo ai fini dell'assunzione di decisioni di investimento. Nel corso degli anni sono state formulate varie teorie per quanto riguarda il processo d'investimento in un portafoglio di strumenti finanziari, ma la prima che diede vita a numerose ricerche in questo settore fu sviluppata da Harry Markowitz a partire dal 1952 e va sotto il nome di Modern Portfolio Theory. Il primo capitolo della tesi è dedicato alla descrizione di tale teoria e delle fasi in cui si dovrebbe strutturare il processo d'investimento. Successivamente vengono presentate alcune critiche alla Modern Portfolio Theory, una delle quali vede la varianza come una misura di rischio non idonea per quantificare la rischiosità di un certo investimento. Il secondo capitolo si propone perciò di illustrare i limiti delle misure di rischio classiche, come la varianza e premoderne, come il Value-at-Risk, sino a giungere all'individuazione di una classe di misure definite coerenti, poiché soddisfano alcune proprietà ritenute fondamentali per un'efficiente ottimizzazione di portafoglio. Viene in seguito illustrata un'innovativa misura di rischio coerente ed entropica che è stata introdotta da Amir Ahmadi-Javid nel 2012 e va sotto il nome di Entropic Value-at-Risk. Tale misura, come si può intuire dal nome, è legata sia al Value-at-Risk che al concetto di entropia relativa ed è la funzione obiettivo del problema di ottimizzazione di portafoglio affrontato nel presente lavoro. Tale funzione, in quanto indicatore della rischiosità del portafoglio, va adeguatamente minimizzata nel rispetto di alcuni vincoli che sono stati illustrati nel terzo capitolo; questi vincoli sono necessari a rendere il problema coerente con la realtà finanziaria. Il problema in questione risulta però essere di notevole complessità, tanto che con le tecniche matematiche classiche non è possibile giungere alla soluzione ottima in un

tempo ragionevole. A fini risolutivi, perciò, verranno usati due algoritmi definiti *metaeuristiche*, che riescono in un breve lasso di tempo a fornire una soluzione sufficientemente vicina a quella ottima; questo è perfettamente in linea con gli interessi di un qualsiasi investitore, poiché, a causa della velocità con cui si evolvono i mercati finanziari, risulta essere sempre meglio assumere buone decisioni d'investimento in tempi brevi che assumere ottime decisioni in tempi decisamente più lunghi. Nel quarto capitolo vengono descritte le metaeuristiche implementate in questo lavoro, ossia la Particle Swarm Optimization, algoritmo che si ispira al comportamento sociale degli stormi di uccelli e dei banchi di pesci e il Fireworks Algorithm, che trae ispirazione dal processo di esplosione dei fuochi d'artificio. Questi due algoritmi, però, nascono per risolvere problemi di ottimizzazione di natura non vincolata e quindi è prima necessario riformulare il problema vincolato in uno senza vincoli. Nel quinto e ultimo capitolo si illustra il metodo delle penalità esatte, necessario per la riformulazione del problema di selezione e infine con Matlab vengono implementati i due algoritmi sui dati reali scaricati dalla piattaforma Bloomberg. I dati sono quelli relativi al FTSE MIB, principale indice azionario italiano e all'OMX Stockholm 30, principale indice azionario svedese. Si vuole dunque ottimizzare due differenti portafogli, uno costituito da azioni italiane e uno formato da azioni svedesi. Vengono, infine, condotti degli studi sia sul funzionamento dei due algoritmi in base al tempo computazionale impiegato e alla bontà delle soluzioni individuate, sia sulla bontà del modello specificato nel presente lavoro. In particolare, suddividendo l'orizzonte temporale di riferimento dei dati storici in due sottoperiodi, quello relativo alla specificazione del modello e quello relativo all'out-of-sample, chiamato anche futuro virtuale, si riesce a testare la bontà del modello sui dati storici, prendendo come riferimento per i test alcuni periodi campionari differenti.

CAPITOLO 1

MODERN PORTFOLIO THEORY

1.1 Aspetti introduttivi

Quando si parla di *investimento* si fa riferimento ad un impiego di denaro o altre risorse a fronte dell'aspettativa di ricavarne un qualche tipo di beneficio in un tempo futuro. L'oggetto di un investimento infatti non è necessariamente il denaro, in quanto a volte noi investiamo il nostro tempo in attività, ad esempio la nostra istruzione. In linea generale si sta rinunciando qualcosa al tempo zero (oggi) sotto l'aspettativa di ottenere qualcosa di maggior valore al tempo uno (domani). Si noti che si parla di aspettativa, in quanto gli eventi futuri non sono conosciuti con certezza. Vi è dunque un'incertezza per quanto riguarda il rendimento futuro e tale aleatorietà dipende dal grado di rischio che ci si sta assumendo in quello specifico investimento.

Tutta la teoria dell'investimento si basa essenzialmente su due concetti essenziali: il *rischio* e il *rendimento*. Essi costituiscono un binomio inscindibile, in quanto devono essere presi congiuntamente in considerazione dagli investitori ai fini della realizzazione di un investimento redditizio. Il legame che sussiste tra rischio e rendimento è diretto, il che significa che ad un aumento della rischiosità di un certo investimento ci si dovrebbe teoricamente aspettare un aumento proporzionale della redditività. Di conseguenza, senza l'assunzione di alcun rischio non ci si dovrebbe aspettare alcun premio derivante dall'investimento medesimo.

Nel corso del presente lavoro si farà esclusivamente riferimento ad una forma di investimento che prevede l'acquisto di strumenti finanziari, come azioni e

obbligazioni, al fine della composizione di un portafoglio con determinate caratteristiche appetibili per il singolo investitore. In questo capitolo verrà illustrata la teoria del portafoglio finanziario, le cui basi risalgono agli anni Cinquanta del secolo scorso e che ha subito profonde innovazioni negli ultimi anni. Questa teoria è di estrema rilevanza in quanto fornisce agli investitori delle linee guida per effettuare degli investimenti coerenti con la loro predisposizione verso il rischio.

Tra i vari studiosi, colui che per primo formalizzò una teoria di portafoglio finalizzata a gestire il rendimento ed il rischio degli investimenti fu Harry Markowitz. Egli, affrontando tematiche strettamente finanziarie, vinse il premio Nobel per l'economia nel 1990. Markowitz divulgò la sua teoria sulla diversificazione e sull'ottimizzazione di portafoglio in un libro dal nome "*Portfolio selection: efficient diversification of investments*"¹, che in parte era già stata espressa in un articolo² del 1952 pubblicato nel *Journal of Finance* in seguito al quale venne poi sviluppata la cosiddetta *Modern Portfolio Theory*.

Markowitz nella sua teoria pionieristica affrontò il tema del rischio in maniera radicalmente differente e si cominciò a parlare non tanto di rischiosità del singolo strumento finanziario, bensì del contributo in termini di rischio che un singolo titolo apporta all'intero portafoglio. Nacque così un rigoroso apparato teorico al fine della costruzione di portafogli efficienti³ di strumenti finanziari, con una rischiosità sufficientemente controllata ed un rendimento conforme alle aspettative del singolo investitore.

Prima di passare all'illustrazione della struttura formale sottostante al processo di ottimizzazione di portafoglio, risulta necessario fornire delle definizioni ai fini della comprensione dei concetti che verranno illustrati in seguito.

¹ H. Markowitz, *Portfolio selection: efficient diversification of investments*, New York, John Wiley & Sons, 1959.

² *Portfolio Selection*, *Journal of Finance*, Vol. 7, No. 1, pp. 77-91, 1952.

³ Il concetto di *efficienza* verrà spiegato nelle pagine successive del presente capitolo.

Data una certa ricchezza iniziale W e indicando con $X = \{X_1, \dots, X_N\}$ l'insieme delle N scelte di investimento, il portafoglio finanziario viene definito come un vettore di dimensioni $1 \times N$:

$$x' = (x_1, \dots, x_N),$$

dove il generico elemento x_i , con $i = 1, \dots, N$, indica la percentuale di W investita in X_i .

Tutta la teoria di portafoglio che verrà descritta nelle seguenti pagine è stata sviluppata partendo dalle seguenti assunzioni (Markowitz, 1952):

- nel mercato non esistono né tasse né costi di transazione⁴ e tutti i titoli sono perfettamente divisibili;
- tutti gli investitori sono *price-taker*⁵ e quindi il loro comportamento non può influenzare la distribuzione di probabilità dei rendimenti degli assets, in quanto il mercato è perfettamente concorrenziale;
- sono ammesse le vendite allo scoperto (*short selling*⁶) senza alcuna restrizione;

⁴ I costi di transazione sono quelli che devono essere sostenuti per realizzare uno scambio, un contratto o in generale una transazione economica. Sono inclusi i costi di ricerca del contraente e quelli relativi ai servizi di intermediazione, i costi per i servizi di consulenti e avvocati e quelli per il monitoraggio dell'attività.

⁵ Un *price-taker* è un operatore di mercato che, date le sue piccole dimensioni in proporzione al mercato in cui esso opera, non ha la possibilità di fissare o influire sui prezzi di beni e servizi. In un mercato perfettamente concorrenziale tutti gli agenti economici sono *price-taker* in quanto i prezzi sono fissati dall'incontro tra domanda e offerta. Un *price-maker*, al contrario, operando in forme di mercato quali monopolio e oligopolio, è in grado di provocare una variazione nel prezzo di mercato di un certo bene, modificando la quantità da lui venduta o acquistata di quel bene.

⁶ La vendita allo scoperto è un'operazione finanziaria che consiste nella vendita di strumenti finanziari non posseduti con successivo riacquisto. Questa operazione si effettua se si ritiene che il prezzo al quale gli strumenti finanziari si riacquisteranno sarà inferiore al prezzo inizialmente incassato attraverso la vendita.

- gli agenti economici conoscono la distribuzione di probabilità dei rendimenti di ogni titolo;
- l'economia è articolata in due soli istanti temporali (t e $t+1$), dove nel primo un individuo sceglie in quali assets investire la propria ricchezza, mentre nel secondo viene misurata la performance, eventualmente corretta per il rischio, di quello specifico investimento, ovvero l'ammontare di perdite o guadagni che si ha conseguito;
- vi è il principio di non sazietà degli investitori poiché, a parità di condizioni, preferiscono sempre un livello di ricchezza maggiore rispetto ad uno minore;
- gli investitori sono avversi al rischio, ovvero tra due titoli con stesso rendimento atteso sceglieranno sempre il titolo con una rischiosità inferiore.

Un investitore che rispetta sia l'assioma di non sazietà sia quello di avversione al rischio è definito razionale.

Le ipotesi appena descritte sono fondamentali ai fini della validità della teoria di portafoglio proposta da Markowitz ma, come si può notare, sono alquanto forti e alcune di esse si discostano nettamente dal mondo reale. Proprio per questi motivi sono state numerose le critiche mosse negli anni a questa impostazione teorica per l'ottimizzazione di portafoglio. Bisogna sottolineare, però, che si sta parlando di modelli finanziari e un modello non deve essere perfettamente consistente con la realtà, deve essere invece una rappresentazione abbastanza fedele ma semplificata della realtà e utile per comprendere determinati aspetti del mondo reale. Va infine specificato che con la Modern Portfolio Theory sono stati introdotti dei modelli che, seppur non esenti da limiti, hanno gettato le basi per tutta la moderna economia finanziaria.

1.2 Il processo di selezione di portafoglio

Il modello di selezione proposto da Markowitz può essere definito come un processo composto da tre fasi consecutive che porta all'identificazione del portafoglio ottimo. Innanzitutto, è necessario identificare degli strumenti per misurare l'incertezza associata a un determinato investimento. Bisogna poi definire un criterio mediante il quale è possibile suddividere tutte le possibili scelte di investimento in un sottoinsieme corrispondente agli investimenti efficienti⁷ e in un altro corrispondente a quelli inefficienti. Infine, si deve scegliere una specifica funzione obiettivo che si vuole ottimizzare⁸, al fine di individuare la scelta di investimento ottimale da detenere tra tutti i possibili investimenti efficienti individuati nel passo precedente.

1.2.1 Misure di rischio e rendimento

Nella prima fase bisogna definire gli strumenti necessari per misurare la redditività ed il rischio di un portafoglio. La maggior parte degli assets finanziari, però, hanno un rendimento aleatorio e di conseguenza risulta impossibile al tempo t conoscere con esattezza il prezzo di un'attività finanziaria al tempo $t+1$. Il rendimento viene perciò modellizzato da una variabile casuale con una certa distribuzione di probabilità.

Markowitz, nel tentativo di catturare l'incertezza associata ad un certo investimento, propose una coppia di misure statistiche legate alla variabile

⁷ Per investimento efficiente si intende un investimento che, a parità di rendimento, ha rischiosità inferiore o specularmente che, a parità di rischio, fornisce il maggior rendimento. Nelle seguenti pagine verrà compiutamente illustrato il criterio per individuare gli investimenti efficienti.

⁸ Si intende sia il processo di minimizzazione del rischio sia quello di massimizzazione del rendimento o di una specifica funzione di utilità.

casuale rendimento: il valore atteso (μ) per misurare la redditività futura e la varianza (σ^2) come misura di rischio di questa redditività.

Partendo da questi concetti di base, i modelli che analizzeremo nel seguente capitolo condividono quindi l'idea che la distribuzione di probabilità dei rendimenti di ogni titolo possa essere interamente descritta dai primi due momenti della distribuzione stessa. Di conseguenza si sta implicitamente assumendo che il rendimento possa anche essere modellizzato da una variabile casuale Normale⁹, in quanto essa è interamente sintetizzata dai parametri valore atteso e varianza.

Considerando una variabile casuale discreta X , che identifica il rendimento aleatorio di un certo strumento finanziario, il valore atteso di tale variabile è il seguente:

$$\mathbb{E}(X) = \sum_{i=1}^N x_i p_i,$$

dove:

- x_i , con $i = 1, \dots, N$, è la i -esima realizzazione di X ;
- p_i , con $i = 1, \dots, N$, è la probabilità che si verifichi l'evento x_i , con $0 \leq p_i \leq 1$ e $\sum_{i=1}^N p_i = 1$.

Nel caso di variabili casuali continue, indicando con $f_X(\cdot)$ la funzione di densità di probabilità, il valore atteso di una generica variabile casuale X è:

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} t f(t) dt;$$

⁹ La distribuzione di probabilità Normale, o gaussiana, è una distribuzione di probabilità continua caratterizzata dal fatto di essere simmetrica rispetto al valore atteso della distribuzione, attorno a cui si concentrano la maggior parte delle osservazioni.

in quest'ultimo caso il valore atteso potrebbe non esistere in caso di divergenza dell'integrale improprio.

Da un punto di vista statistico, la media o valore atteso di una variabile casuale è un indice di posizione; tuttavia da un punto di vista finanziario tale indice viene usato per misurare la profittabilità di una scelta d'investimento.

La varianza di una variabile aleatoria è invece un indice statistico di variabilità. Nel caso di una variabile casuale discreta essa è data dalla media ponderata¹⁰ del quadrato degli scostamenti delle singole realizzazioni dalla media:

$$Var(X) = \sum_{i=1}^N [x_i - \mathbb{E}(X)]^2 p_i.$$

Nel caso di variabili casuali continue, invece, indicando con $f_X(\cdot)$ la funzione di densità di probabilità, la formula della varianza è la seguente:

$$Var(X) = \int_{-\infty}^{+\infty} [t - \mathbb{E}(X)]^2 f(t) dt.$$

La varianza del tasso di rendimento, da un punto di vista finanziario, è una misura di rischio di una scelta d'investimento. Essa cattura infatti la variabilità del rendimento stesso, più o meno accentuata a seconda di quanto sono distanti le singole osservazioni dal rendimento medio.

Anche la varianza, nel caso di variabili casuali continue, può non esistere in caso di divergenza dell'integrale improprio.

¹⁰ I pesi della ponderazione sono dati dalle probabilità che si verifichino le singole realizzazioni.

A questo punto è utile illustrare come determinare il rendimento atteso e la varianza nel caso non di un singolo strumento finanziario, ma di un portafoglio di titoli. Il rendimento di portafoglio R_p è semplicemente dato dalla media ponderata dei rendimenti dei singoli titoli che lo compongono:

$$R_p = \sum_{i=1}^N x_i R_i.$$

Applicando alcune semplici proprietà del valor medio, la media del rendimento di portafoglio viene definita come:

$$\mathbb{E}(R_p) = \sum_{i=1}^N x_i r_i \stackrel{\text{def}}{=} r_p,$$

con notazione vettoriale $r_p = \mathbf{x}'\mathbf{r}$,

dove r_i è il rendimento medio del titolo i -esimo. Da tale formalizzazione si evince che per costruire un portafoglio a rendimento atteso più elevato è necessario investire un ammontare maggiore di ricchezza nei titoli con un rendimento medio maggiore.

Il rendimento però da solo non riesce a fornire le informazioni cruciali in merito alla bontà di un investimento. Non bisogna mai dimenticare, infatti, il binomio inscindibile tra rischio e rendimento. Il rischio, che nella Modern Portfolio Theory viene misurato dalla varianza, riveste un'importanza chiave in quanto fornisce un'indicazione della rischiosità o volatilità dell'investimento. Un qualunque investitore razionale, se dovesse scegliere tra due titoli caratterizzati da una diversa varianza ma medesimo rendimento atteso, preferirebbe investire nel titolo a rischiosità inferiore; questo è coerente con il principio di avversione al rischio di cui si è accennato poco sopra. Il processo di misurazione e gestione

del rischio, che verrà illustrato di seguito, risulta estremamente rilevante ai fini della corretta ideazione di un processo di investimento.

La varianza di portafoglio è più complessa da calcolare rispetto al rendimento atteso, in quanto è pari alla media ponderata delle varianze dei singoli titoli soltanto se tali titoli sono tra loro stocasticamente indipendenti¹¹. Tuttavia nei mercati finanziari la maggior parte delle azioni presenta un certo livello di dipendenza, specie se si tratta di società appartenenti allo stesso settore o a settori affini. Per questi motivi nella formula della varianza di portafoglio è necessario tenere in considerazione la dipendenza che può sussistere tra coppie di rendimenti associati a titoli diversi. Si usa pertanto la covarianza, che è un indicatore statistico in grado di misurare la dipendenza lineare tra le coppie di titoli che compongono il portafoglio. La varianza del rendimento di portafoglio è quindi la seguente:

$$\text{Var}(R_p) = \sum_{i=1}^N x_i^2 \sigma_i^2 + 2 \sum_{i=1}^N \sum_{j=i+1}^N x_i x_j \sigma_{ij} \stackrel{\text{def}}{=} \sigma_p^2,$$

dove σ_{ij} è la covarianza tra il titolo *i*-esimo e il titolo *j*-esimo.

Come si può notare, la varianza di portafoglio è in parte determinata dalle varianze dei rendimenti dei singoli titoli e in parte dalle covarianze esistenti tra tutte le coppie di titoli. Può inoltre essere scritta in forma matriciale nel seguente modo:

$$\sigma_p^2 = \mathbf{x}' \mathbf{V} \mathbf{x},$$

¹¹ L'indipendenza stocastica ha un'accezione più ampia rispetto all'indipendenza lineare (misurata dall'indice di correlazione di Pearson). Si parla generalmente di indipendenza stocastica tra due eventi A e B quando il verificarsi dell'uno non modifica la probabilità di verificarsi dell'altro.

dove x è il vettore colonna delle percentuali investite nei singoli titoli, ' indica l'operazione di trasposizione e V è la matrice di varianze e covarianze dei rendimenti¹². Il risultato di tale prodotto matriciale non può essere negativo, in quanto per definizione la varianza è sempre maggiore o uguale a zero. La covarianza, invece, ha un campo di variazione che va da $-\infty$ a $+\infty$ e di conseguenza risulta assai complesso l'utilizzo di tale misura a fini comparativi. Per risolvere questo problema di comparabilità si ricorre al coefficiente di correlazione di Pearson, ottenuto dividendo la covarianza per il prodotto delle deviazioni standard¹³, che presenta il vantaggio di variare all'interno dell'intervallo chiuso $[-1;1]$:

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

Il coefficiente di correlazione misura l'eventuale legame lineare sussistente tra due variabili, in questo caso tra i rendimenti di due titoli. Un indice pari a +1 indica una perfetta correlazione positiva tra due titoli, il che implica che un aumento percentuale del rendimento dell'uno comporta un aumento percentuale dell'altro secondo una relazione lineare. La situazione si capovolge se l'indice è pari a -1, ossia nel caso di perfetta correlazione negativa, mentre nel caso in cui sia pari a 0 vi è una totale assenza di correlazione e quindi non vi è alcun legame lineare tra i titoli in questione.

Illustrato il significato del coefficiente di correlazione, possiamo ora riscrivere la formula della varianza di portafoglio nella maniera seguente:

¹² È una matrice simmetrica in cui le varianze dei singoli titoli compaiono sulla diagonale principale e le covarianze al di fuori della diagonale stessa.

¹³ La deviazione standard, o scarto quadratico medio, è la radice quadrata della varianza.

$$\sigma_p^2 = \sum_{i=1}^N x_i^2 \sigma_i^2 + 2 \sum_{i=1}^N \sum_{j=i+1}^N x_i x_j \rho_{ij} \sigma_i \sigma_j.$$

Il più grande vantaggio di investire in un insieme di titoli piuttosto che nei titoli singolarmente è quello legato alla contrazione del rischio, fenomeno strettamente connesso alla cosiddetta *diversificazione*. Aggregando strumenti finanziari diversi si può riuscire a creare uno strumento composto con una rischiosità inferiore a quella dei singoli titoli che lo compongono. L'effetto della diversificazione dipende dalle correlazioni esistenti tra i titoli all'interno del portafoglio: più le correlazioni sono basse o, meglio ancora, negative più la diversificazione contribuisce ad una maggiore riduzione del rischio complessivo, in quanto i rendimenti dei titoli tendono a muoversi in direzioni opposte, creando un effetto di compensazione tra profitti e perdite e riducendo così la volatilità complessiva del portafoglio. Se il processo di selezione di portafoglio viene condotto efficacemente, quindi, la diversificazione consente di ottenere discreti benefici in termini di rischiosità complessiva.

Considerando la fattispecie dei titoli azionari, si può dunque affermare che l'inserimento di azioni con covarianze basse all'interno di un portafoglio comporta una contrazione del rischio complessivo. Per comprendere ciò che avviene esattamente bisogna scomporre il rischio in due fattispecie: il rischio *diversificabile* e quello *non diversificabile*.

Il rischio diversificabile, o specifico, attiene alle caratteristiche della singola impresa, ed è misurato dalla varianza di un determinato titolo che definisce la rischiosità specifica di quel business. È una fonte di rischio che si può ridurre ed eliminare attraverso la diversificazione, nei modi che sono stati descritti precedentemente.

Il rischio non diversificabile, chiamato anche rischio di mercato, è invece misurato dalle covarianze tra le coppie di titoli ed è legato all'andamento dell'intero mercato di riferimento. Esso non può essere ridotto diversificando il

portafoglio, in quanto si tratta di dinamiche che non riguardano il singolo titolo, bensì che riguardano l'insieme di titoli facenti parte, ad esempio, di un certo settore o indice azionario.

La covarianza, e di conseguenza anche la correlazione, riesce a cogliere soltanto le dipendenze lineari tra due variabili casuali. Questo indicatore statistico non considera gli eventuali legami non lineari che possono sussistere tra i rendimenti di due strumenti finanziari. Questa peculiarità rende la covarianza una misura di dipendenza parziale, in quanto si può affermare che l'indipendenza tra due titoli implichi una covarianza - o correlazione - nulla, mentre una covarianza nulla non implica una totale indipendenza tra le due variabili casuali che modellizzano i rendimenti dei titoli, questo perché potrebbero sussistere dei legami di natura non lineare che il coefficiente di correlazione non riesce ad individuare.

I mercati finanziari odierni sono spesso caratterizzati da dinamiche estremamente complesse e di conseguenza si può ben comprendere come l'approccio media-varianza e l'indice di correlazione lineare non riescano a cogliere completamente le dinamiche di mercato; tale limite è tanto più evidente quanto è maggiore la complessità dei mercati in cui si sta operando.

1.2.2 Criterio di dominanza stocastica e frontiera efficiente

Illustrati gli strumenti necessari per misurare l'aleatorietà di un investimento, si può passare alla seconda fase del processo di ottimizzazione di portafoglio. Bisogna ora, sulla base dei concetti di rendimento atteso e varianza di cui si è parlato poco sopra, suddividere tutte le possibili scelte di investimento in due gruppi. I due gruppi sono mutualmente esclusivi e uno contiene tutti gli investimenti efficienti mentre l'altro tutti quelli inefficienti. In questo modo risulta già possibile una prima scrematura delle alternative su cui si potrebbe

investire. Questa fase, quindi, come affermato dal matematico Szegö¹⁴, si traduce nell'identificazione di un criterio di efficienza, mediante il quale distinguere quali sono le alternative da scartare e quali quelle da valutare nella fase successiva ai fini della costruzione del portafoglio ottimo.

Il criterio a cui si fa riferimento viene definito *criterio di dominanza stocastica*: basandosi sull'approccio media-varianza¹⁵ si vuole così suddividere i titoli efficienti da quelli inefficienti, ossia rispettivamente quelli che dominano altri titoli da quelli che vengono dominati, secondo delle regole ben precise.

Consideriamo il caso più generale in cui X_1 e X_2 siano due variabili casuali che, ad esempio, modellizzino i rendimenti di due titoli azionari; si afferma che X_1 domina X_2 , oppure che X_1 è preferito a X_2 secondo il criterio media-varianza se:

$$\mathbb{E}(X_1) \geq \mathbb{E}(X_2) \quad \text{e} \quad \text{Var}(X_1) < \text{Var}(X_2)$$

oppure se

$$\mathbb{E}(X_1) > \mathbb{E}(X_2) \quad \text{e} \quad \text{Var}(X_1) \leq \text{Var}(X_2).$$

Infatti a parità di rendimento atteso, un investitore razionale preferisce un titolo con una rischiosità inferiore e a parità di varianza viene preferito il titolo che in media presenta rendimenti maggiori. Si ha indifferenza tra due strumenti finanziari quando essi presentano stessi valori di varianza e rendimento atteso.

Dal punto di vista della notazione possiamo scrivere quanto segue:

- $X_1 \succ_{MV} X_2$: X_1 domina in media-varianza X_2 ;
- $X_1 \sim_{MV} X_2$: X_1 è indifferente a X_2 secondo il criterio media-varianza;
- $X_1 \succcurlyeq_{MV} X_2$: X_1 domina o è indifferente in media-varianza a X_2 .

¹⁴ Giorgio P. Szegö, *Portfolio Theory: With Application to Bank Asset Management*, New York, Academy Press, 1980 (pagina 8).

¹⁵ Tale approccio è illustrato ed esemplificato qui di seguito.

Il criterio di dominanza stocastica appena presentato ha però un limite: in alcuni casi esso non riesce a stabilire un ordine di preferenza e quindi non risulta un criterio di scelta che soddisfa la completezza¹⁶. Si sta facendo riferimento, ad esempio, al caso in cui un titolo presenti sia la varianza sia il rendimento atteso maggiori rispetto a quelli di un altro titolo.

Di seguito si riporta un esempio dell'applicazione del criterio media-varianza a tre portafogli A, B, C. Il portafoglio A presenta un rendimento medio del 10% e una deviazione standard¹⁷ del 3%, B ha un rendimento medio del 5% e una deviazione standard del 4%, mentre C ha un rendimento medio del 4% e una deviazione standard del 3,5%. Come si può notare dalla Figura 1.1, il portafoglio A domina in media-varianza sia B che C, dato che esso presenta sia un rendimento atteso più elevato che una deviazione standard inferiore agli altri due portafogli.

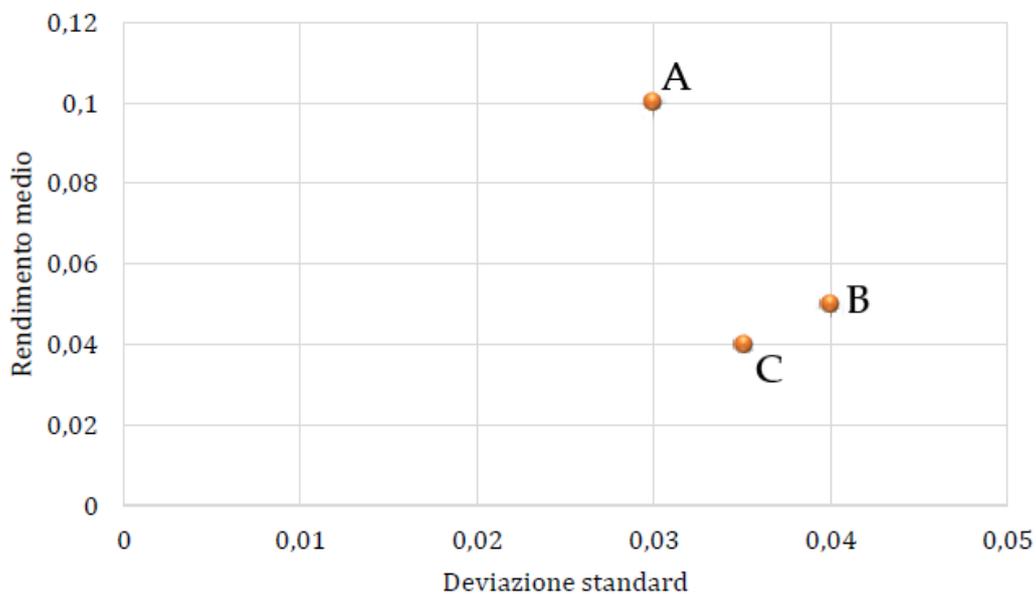


Figura 1.1 - Esempificazione del criterio media-varianza

¹⁶ La completezza è la proprietà per cui un individuo sia sempre in grado di scegliere tra due alternative.

¹⁷ Nella pratica si usa la deviazione standard, che è la radice quadrata della varianza, come indicatore di volatilità dei rendimenti e quindi di rischio. Si noti che la deviazione standard ha la stessa unità di misura della media.

Come è stato detto prima, però, questo criterio di scelta non sempre porta ad un ordinamento completo di titoli o portafogli. Con l'approccio media-varianza non è infatti possibile stabilire un ordine di preferenza tra i portafogli B e C, in quanto B è caratterizzato sia da un rendimento atteso che da una deviazione standard maggiori rispetto a C.

Tutti i portafogli classificati come efficienti formano la cosiddetta *frontiera efficiente*, che risulta essere l'arco superiore di una parabola nel piano media-varianza e l'arco superiore di un'iperbole nel piano media-deviazione standard. Tra tutti i portafogli possibili che si possono costruire con un determinato numero di titoli, tale frontiera comprende quelli per cui, dato un certo rendimento atteso, è minimizzata la varianza oppure simmetricamente quelli per cui, data una certa varianza, è massimizzato il rendimento atteso. La scelta di un individuo razionale dovrà perciò ricadere su uno dei portafogli facenti parte di questa frontiera, in quanto efficienti.

La Modern Portfolio Theory introdotta da Markowitz fornisce un rigoroso criterio per determinare questa frontiera efficiente. È necessario impostare un problema di programmazione quadratica come il seguente:

$$\begin{aligned} & \min_{x_1, \dots, x_N} \sigma_p^2 \\ & \text{s. t.} \quad \begin{cases} \mathbb{E}(r_p) = \pi \\ \sum_{i=1}^N x_i = 1 \end{cases} \end{aligned}$$

che può anche essere scritto in forma matriciale:

$$\begin{aligned} & \min_{x_1, \dots, x_N} \mathbf{x}'\mathbf{V}\mathbf{x} \\ & \text{s. t.} \quad \begin{cases} \mathbf{x}'\mathbf{r} = \pi \\ \mathbf{x}'\mathbf{e} = 1 \end{cases} \end{aligned}$$

dove:

- π è il valore del rendimento atteso desiderato dall'investitore;
- x rappresenta il vettore contenente le percentuali incognite di capitale da investire negli N titoli;
- V è la matrice $N \times N$ di varianze e covarianze;
- r è il vettore dei rendimenti attesi degli N titoli;
- e è un vettore unitario di dimensione N.

Quanto descritto poco sopra rappresenta in dettaglio un problema di minimizzazione vincolata, in quanto viene minimizzata la funzione obiettivo - che misura la rischiosità del portafoglio - nel rispetto di due vincoli lineari. In particolare il primo vincolo definisce il rendimento atteso che si vuole ottenere dal portafoglio in questione e il secondo invece, chiamato *vincolo di bilancio*, pone la condizione che tutto il capitale disponibile sia investito.

Risolvendo questo problema di programmazione si trovano le coordinate varianza-rendimento atteso di ogni portafoglio efficiente per ciascun livello di rendimento desiderato e viene così costruita la frontiera efficiente (ed anche quella inefficiente). Nella Figura 1.2 viene esemplificato quanto appena detto e si può notare che i portafogli appartenenti alla frontiera appena illustrata sono quelli con il rischio minimo per ogni livello di rendimento atteso.

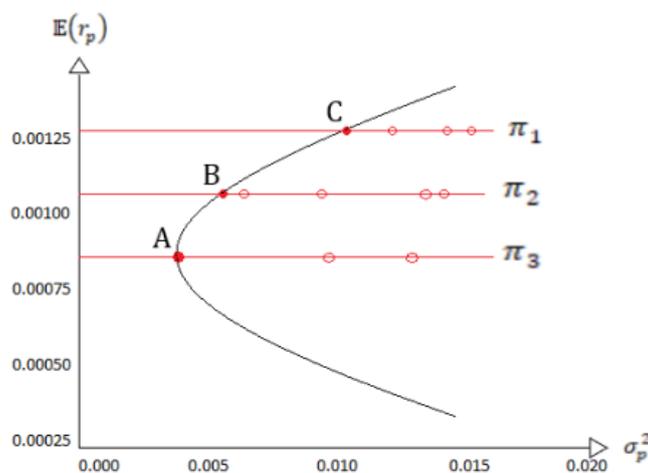


Figura 1.2 - Esempio di frontiera efficiente e inefficiente

Nella figura in questione la parabola rappresenta la frontiera di tutti i possibili portafogli a varianza minima che si possono costruire; questi portafogli, infatti, sono chiamati *minimum variance portfolios*.

Il portafoglio A ha in assoluto la varianza e quindi la rischiosità inferiore e va sotto il nome di *global minimum variance portfolio*. Questo portafoglio suddivide l'intera curva in due regioni: quella inferiore comprende tutte le combinazioni inefficienti di titoli e quella superiore, che è la frontiera efficiente, comprende tutti i cosiddetti *efficient minimum variance portfolios*, tra cui ad esempio i portafogli A, B e C. I portafogli che stanno sulla parte inferiore della parabola sono inefficienti in quanto per ciascuno di essi esiste almeno un altro portafoglio con la stessa varianza ma con un rendimento atteso maggiore, risultando così dominante in media-varianza.

Come conseguenza dell'applicazione di questo criterio di dominanza stocastica, la frontiera su cui giacciono tutti i portafogli efficienti risulta essere sempre una funzione concava.

Per risolvere il problema di minimizzazione della varianza proposto precedentemente, bisogna innanzitutto definire le ipotesi che devono essere soddisfatte affinché il problema di ottimizzazione abbia una soluzione univoca, rappresentata dal vettore contenente le percentuali ottime di capitale investito in ciascun titolo del portafoglio.

La prima ipotesi è che la matrice di varianze e covarianze sia non singolare, ossia che tutte le righe e le colonne di tale matrice siano linearmente indipendenti. In termini finanziari, questo significa che non devono esistere titoli con strutture di rischio esprimibili come combinazioni lineari delle strutture di rischio di altri titoli.

La seconda condizione che dev'essere soddisfatta è che la matrice di varianze e covarianze sia definita positiva. Ne consegue che tutte le varianze dei singoli

titoli sono positive. Non vengono quindi considerati i titoli risk-free, ossia quelli a rendimento certo e, quindi, varianza nulla.

Infine, l'ultima condizione, è che esistano almeno due titoli con rendimenti attesi diversi, altrimenti il problema di selezione diventerebbe banale, in quanto tutto il capitale verrebbe investito esclusivamente nel titolo con varianza inferiore.

Una volta che queste ipotesi sono rispettate, allora si può dimostrare che la varianza del rendimento del portafoglio è una funzione convessa e in quanto tale esiste un solo punto di minimo assoluto. Si ha così esistenza ed unicità della soluzione.

Per calcolare la soluzione di questo problema di minimizzazione bisogna innanzitutto costruire la seguente funzione Lagrangiana:

$$L = \mathbf{x}'\mathbf{V}\mathbf{x} - \lambda_1(\mathbf{x}'\mathbf{r} - \pi) - \lambda_2(\mathbf{x}'\mathbf{e} - 1).$$

Si imposta poi un sistema di tre condizioni, le cosiddette first-order conditions (FOC¹⁸), differenziando la Lagrangiana rispetto a ogni variabile e rispetto ai moltiplicatori di Lagrange (λ_1 e λ_2) e ponendo le equazioni risultanti pari a zero:

$$\begin{cases} \frac{\partial L}{\partial \mathbf{x}} = 2\mathbf{x}'\mathbf{V} - \lambda_1\mathbf{r}' - \lambda_2\mathbf{e}' = \mathbf{0}_N \\ \frac{\partial L}{\partial \lambda_1} = -\mathbf{x}'\mathbf{r} + \pi = 0 \\ \frac{\partial L}{\partial \lambda_2} = -\mathbf{x}'\mathbf{e} + 1 = 0 \end{cases}$$

¹⁸ Le condizioni del primo ordine sono le condizioni necessarie che devono essere soddisfatte da ogni soluzione di un problema di ottimizzazione, ossia in ogni punto di massimo o minimo. Un punto che soddisfa tali condizioni si dice anche punto critico, perché candidato a essere una soluzione del problema di ottimo.

Dopo una serie di passaggi algebrici, si arriva alla seguente soluzione:

$$\mathbf{x}^* = \frac{(\gamma \mathbf{V}^{-1} \mathbf{r} - \beta \mathbf{V}^{-1} \mathbf{e})\pi + (\alpha \mathbf{V}^{-1} \mathbf{e} - \beta \mathbf{V}^{-1} \mathbf{r})}{\alpha\gamma - \beta^2}$$

dove:

- \mathbf{x}^* è il vettore contenente le percentuali di ricchezza investita nei singoli titoli che compongono il portafoglio ottimo selezionato;
- $\alpha = \mathbf{r}'\mathbf{V}^{-1}\mathbf{r}$;
- $\beta = \mathbf{r}'\mathbf{V}^{-1}\mathbf{e} = \mathbf{e}'\mathbf{V}^{-1}\mathbf{r}$;
- $\gamma = \mathbf{e}'\mathbf{V}^{-1}\mathbf{e}$.

Questa formula consente di determinare tutti i portafogli efficienti per ogni valore del *target return*, ossia del rendimento desiderato π . È così possibile determinare la frontiera efficiente.

Riformulando i risultati precedenti, l'espressione della frontiera può essere espressa in termini di varianza come:

$$\sigma_{p^*}^2 = \mathbf{x}^{*\prime} \mathbf{V} \mathbf{x}^* = \frac{\gamma\pi^2 - 2\beta\pi + \alpha}{\alpha\gamma - \beta^2}$$

e in termini di deviazione standard come:

$$\sigma_{p^*} = (\mathbf{x}^{*\prime} \mathbf{V} \mathbf{x}^*)^{1/2} = \left(\frac{\gamma\pi^2 - 2\beta\pi + \alpha}{\alpha\gamma - \beta^2} \right)^{1/2}.$$

Date queste due formule, si può derivare la forma della frontiera nel piano media-varianza e nel piano media-deviazione standard: nel primo caso è una parabola dove il vertice ha coordinate $\left(\sigma_{p,v}^2 = \frac{1}{\gamma}, r_{p,v} = \frac{\beta}{\gamma} \right)$, mentre nel secondo

caso è un'iperbole dove il vertice ha coordinate $\left(\sigma_{p,v} = \left(\frac{1}{\gamma}\right)^{\frac{1}{2}}, r_{p,v} = \frac{\beta}{\gamma}\right)$.

L'identificazione del portafoglio posizionato nel vertice è cruciale in quanto esso divide la frontiera dei possibili portafogli in una parte inferiore, la cosiddetta *frontiera inefficiente* ed in una parte superiore, la *frontiera efficiente*. Il portafoglio situato nel vertice è quello che fra tutti i portafogli efficienti ha la minore varianza e quindi una rischiosità più contenuta, che lo rende un'interessante scelta di investimento per i periodi di crisi finanziaria.

Nella Figura 1.3 si possono osservare due frontiere efficienti ottenute con Matlab, sia nel piano media-varianza sia in quello che considera la deviazione standard come indice di rischiosità. Per la loro rappresentazione si è fatto riferimento ai prezzi di chiusura giornalieri¹⁹ di sei titoli azionari²⁰ che, assieme ad altri, compongono l'indice FTSE MIB. Quest'ultimo comprende le azioni delle quaranta società italiane quotate con maggiore capitalizzazione²¹, flottante²² e liquidità²³.

La frontiera efficiente è rappresentata dal lato superiore della linea continua, divisa in due sottoinsiemi dal portafoglio raffigurato dal pallino nero, che è quello a rischiosità inferiore tra tutti quelli teoricamente costruibili. Tutti i portafogli situati al di sotto della frontiera efficiente sono inefficienti, in quanto risultano essere dominati in media-varianza da almeno un altro portafoglio. I

¹⁹ Si tratta di 252 osservazioni, corrispondenti ad altrettanti giorni di contrattazione.

²⁰ I titoli presi in considerazione sono UniCredit, Telecom Italia, Moncler, Intesa Sanpaolo, Generali ed Enel.

²¹ La capitalizzazione di mercato è il prodotto tra il numero di azioni di una società e il prezzo di mercato della singola azione.

²² Il flottante è la quantità di azioni che gli investitori possono liberamente commerciare nel mercato secondario. Normalmente corrisponde ad una percentuale del totale delle azioni emesse da una determinata società.

²³ Per liquidità si intende l'attitudine di un investimento a trasformarsi in denaro velocemente e possibilmente senza perdite. Normalmente uno strumento tanto più è liquido tanto più è caratterizzato da una grande quantità di contrattazioni di borsa e quindi da un forte interesse da parte degli operatori.

pallini rossi rappresentano invece la varianza (o deviazione standard) e il rendimento atteso dei sei titoli presi singolarmente. Si noti come è possibile ridurre significativamente il rischio investendo in un paniere di appena sei titoli; questo è appunto l'effetto della diversificazione, che è tanto maggiore quanto più basse sono le correlazioni lineari tra le coppie di assets. Dalla Figura 1.3 risultano perciò chiari i benefici che si possono ottenere con la composizione adeguata di un portafoglio, in termini di rischio e di rendimento atteso. Come sottolineato precedentemente, infine, una volta trovata la frontiera efficiente, la scelta d'investimento del singolo investitore dipenderà dal suo grado di avversione al rischio, parametro soggettivo di ogni individuo.

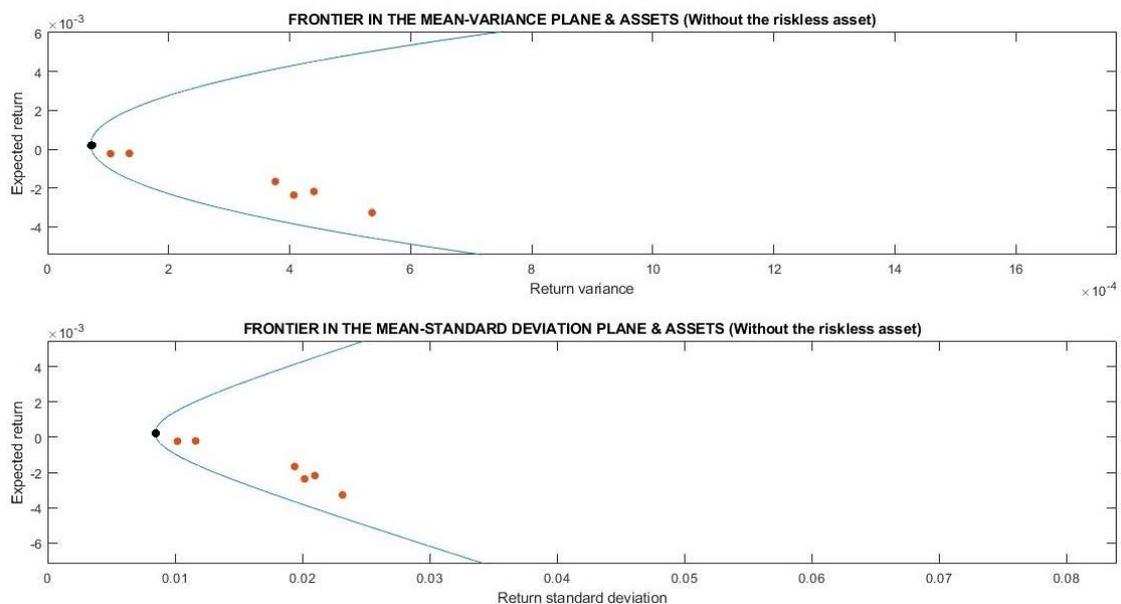


Figura 1.3 - Frontiera efficiente e inefficiente nel piano media-varianza e media-deviazione standard con $N=6$

Alla base di tutto il processo di composizione e ottimizzazione di portafoglio c'è il concetto di diversificazione, di cui si è parlato nelle pagine precedenti. A tal proposito, per chiarire questo concetto collegato al tema della contrazione del rischio del portafoglio rispetto a quello dei vari singoli assets che lo compongono, risulta utile prendere in considerazione una particolare strategia

d'investimento²⁴ che porta alla costruzione del portafoglio *equally-weighted*. Questo portafoglio è equipesato, in quanto si investe la stessa percentuale di capitale in ogni titolo componente il portafoglio stesso; in altri termini ogni singolo titolo ha lo stesso peso percentuale all'interno del portafoglio.

La varianza dei rendimenti di tale portafoglio può essere calcolata come:

$$\sigma_p^2 = \frac{1}{N} \sum_{i=1}^N \frac{\sigma_i^2}{N} + \frac{N-1}{N} \sum_{i=1}^N \sum_{j=1, j \neq i}^N \frac{\sigma_{ij}}{N(N-1)}.$$

Si noti che la prima sommatoria rappresenta la varianza media dei rendimenti dei titoli componenti il portafoglio, mentre la successiva doppia sommatoria indica la covarianza media tra i rendimenti degli stessi. Tale formulazione può perciò essere riscritta come segue:

$$\sigma_p^2 = \frac{1}{N} \bar{\sigma}_i^2(N) + \frac{N-1}{N} \bar{\sigma}_{ij}(N).$$

Quando N tende a infinito il primo addendo tende a zero, mentre il secondo addendo tende a $\bar{\sigma}_{ij}(N)$ dato che il rapporto $\frac{N-1}{N}$ tende a uno.

Questa formulazione rende evidente il significato della diversificazione, secondo cui il rischio di portafoglio diminuisce all'aumentare del numero dei titoli che lo compongono. Il rischio però non è totalmente eliminabile, in quanto se il numero dei titoli è sufficientemente elevato la varianza di portafoglio tende alla covarianza media dei titoli. Con la diversificazione si riesce quindi ad eliminare il *rischio idiosincratico*, che è specifico dei singoli titoli, mentre rimane una componente di rischiosità non diversificabile, definita dalle covarianze e chiamata *rischio sistematico* o *rischio di mercato*. La Figura 1.4 esemplifica quanto

²⁴ La strategia in questione è chiamata *strategia 1/N*, perché il capitale disponibile viene suddiviso equamente tra gli N titoli componenti il portafoglio.

appena detto, evidenziando una relazione inversa tra la varianza di portafoglio e il numero di titoli che lo compongono.

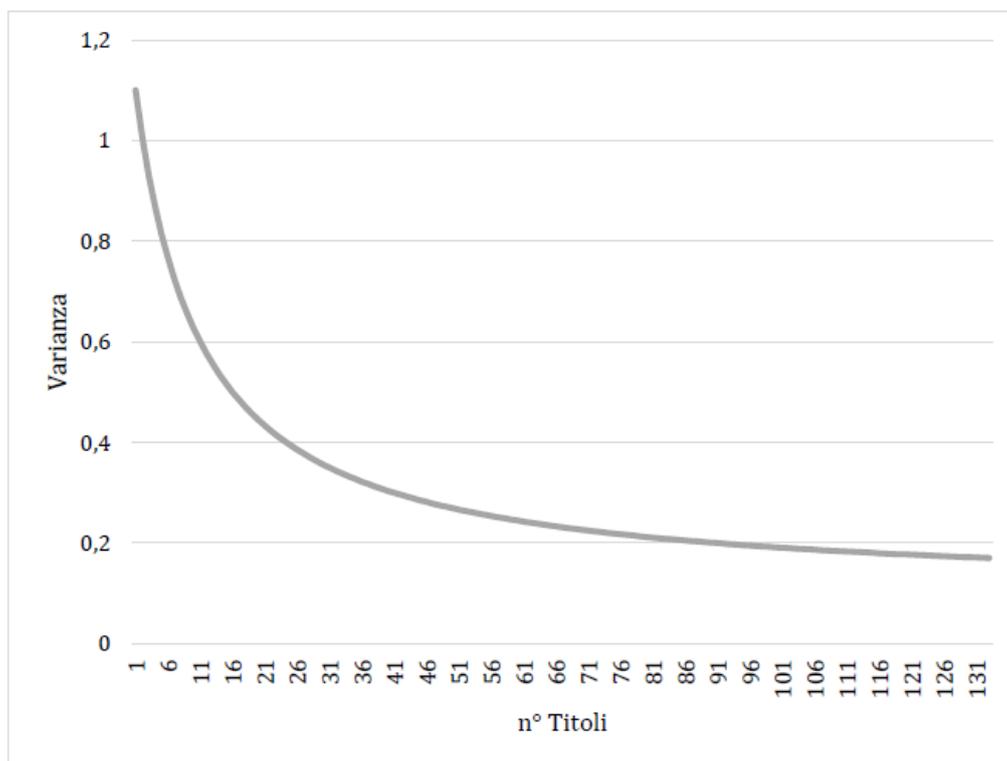


Figura 1.4 - Relazione tra varianza di portafoglio e numero di titoli che lo compongono

Si osservi infine che gli effetti della diversificazione sono tanto più accentuati quanto più la numerosità di titoli è bassa. È perciò sufficiente inserire in portafoglio un numero ridotto di titoli per beneficiare già di una contrazione del rischio. Questi concetti sono in linea con la realtà dei mercati finanziari. Il numero di scelte di investimento è infatti limitato e i costi da sostenere se si volesse investire in un gran quantitativo di titoli sarebbero troppo elevati per gli investitori. Per queste ragioni è necessario costruire i portafogli di strumenti finanziari in modo da beneficiare degli effetti della diversificazione senza incorrere in eccessivi costi di acquisto relativi agli strumenti stessi.

1.2.3 Dalla teoria dell'utilità attesa al portafoglio ottimo

La terza e ultima fase del processo di selezione di portafoglio ideato da Markowitz consiste nel selezionare, tra tutti i portafogli appartenenti alla frontiera efficiente, quello che meglio soddisfa il profilo di rischio-rendimento del singolo investitore. Per comprendere appieno quest'ultima fase è innanzitutto necessario introdurre sinteticamente la teoria dell'utilità attesa.

La teoria dell'utilità attesa è stata sviluppata da Von Neumann e Morgenstern nel 1947²⁵ ed è un modello ampiamente utilizzato per spiegare le scelte degli individui in condizioni di incertezza. Questa teoria considera l'uomo come un essere razionale. Quando un individuo si trova di fronte a delle alternative senza conoscere con certezza quale si verificherà, ma ne conosce la distribuzione di probabilità di realizzazione, secondo la teoria egli sceglie quella per cui si aspetta un'utilità attesa più elevata. Con il termine *utilità* si può identificare il grado di soddisfacimento di un certo individuo e quindi la massimizzazione dell'utilità attesa è il fine ultimo di ciascun essere razionale, sia consumatore che investitore.

A questo punto sorge inevitabile il confronto con la teoria microeconomica del consumatore. La funzione di utilità è qui una funzione che tiene conto della quantità di beni e servizi consumati e non tiene invece in considerazione la probabilità di accadimento degli eventi, per il fatto che la teoria del consumo si basa essenzialmente su eventi certi. Inoltre l'utilità del consumatore ha un massimo oltre il quale non può più crescere, perché le risorse a disposizione degli individui sono scarse. Parlando di scarsità delle risorse si fa riferimento sia al denaro, indispensabile per l'acquisto di beni e servizi, sia all'ammontare disponibile dei beni e servizi stessi.

²⁵ J. Von Neumann & O. Morgenstern, *Theory of games and economic behaviour*, Princeton, New Jersey, US, Princeton University Press, 1947.

La teoria dell'investimento, che è alla base del presente lavoro, è legata alla teoria del consumatore ma la funzione di utilità in questione è diversa. Affrontando il tema dell'incertezza risulta necessario l'utilizzo di un'altra funzione di utilità, che tenga conto dell'aleatorietà degli eventi, intendendo in questo caso il rischio legato ai rendimenti dei titoli. Questa funzione è una *funzione di utilità della moneta*, in quanto tiene conto del soddisfacimento che può portare ad un certo investitore il guadagno di un determinato ammontare certo di denaro. Essa è crescente²⁶, anche se l'utilità marginale derivante da un incremento della ricchezza diminuisce all'aumentare della ricchezza posseduta dall'individuo. L'utilità marginale decrescente recepisce il fatto che gli individui sono avversi al rischio e questo implica la concavità della funzione di utilità. La maggior parte degli individui di fatto presenta un certo grado di avversità al rischio ma vi sono agenti non razionali che sono predisposti al rischio oppure totalmente indifferenti, e la funzione di utilità risulta essere convessa per i primi e lineare invece per i secondi.

Parlando di avversione al rischio si fa riferimento ad un atteggiamento psicologico che caratterizza la maggior parte delle persone e che le rende particolarmente attente al profilo di rischio di un certo investimento. Tali soggetti sono disposti ad accettare rendimenti attesi assumendo un rischio il più contenuto possibile. Misurare l'avversione al rischio di un individuo è complesso, perché dipende da una moltitudine di fattori soggettivi come ad esempio il livello personale di percezione del rischio, la situazione finanziaria, l'età e la prospettiva di reddito futuro. A tal proposito viene normalmente utilizzato un questionario che restituisce una valutazione qualitativa dell'individuo²⁷.

²⁶ Il benessere degli individui aumenta all'aumentare della loro ricchezza (principio di non sazietà).

²⁷ N. Linciano, P. Soccorso, *La rilevazione della tolleranza al rischio degli investitori attraverso il questionario*, Discussion Papers, CONSOB, 2012.

Tornando al nostro problema di partenza, per stabilire quale tra i portafogli appartenenti alla frontiera efficiente risulti essere quello che meglio si adatta alle proprie esigenze e propensioni al rischio, è necessario una funzione di utilità che tenga conto della soggettività del singolo investitore. A tale fine risulta cruciale la teoria dell'utilità attesa.

L'utilità attesa del rendimento di un portafoglio, come ci si può aspettare, è una funzione del rendimento medio e della varianza del rendimento del portafoglio. Posto alla base il principio di razionalità degli individui, vi è una relazione crescente con il rendimento atteso ed inversa con la varianza. Quindi, dato il fine di massimizzare l'utilità attesa, gli investitori, a seconda del loro grado di avversione al rischio, possono servirsi di tale funzione per scegliere quale tra gli investimenti efficienti è più consono alle proprie esigenze.

Dopo queste prime considerazioni risulta necessario fare una distinzione tra *utilità degli importi certi* e *utilità attesa degli importi aleatori*. Per quanto riguarda la prima bisogna sottolineare che non vi è alcuna componente aleatoria e un individuo sceglierà l'investimento x_1 rispetto a x_2 se la sua utilità, legata al rendimento certo che tale investimento produrrà, è superiore all'utilità di x_2 :

$$u(x_1) > u(x_2) \Leftrightarrow x_1 > x_2.$$

Questa relazione mette anche in luce la crescita della funzione di utilità di cui si è parlato poco sopra. La concavità, e quindi l'avversione al rischio degli investitori, dati $x_1 < x_2$ e una variazione di importo monetario Δx , viene invece specificata dalla seguente relazione:

$$u(x_1 + \Delta x) - u(x_1) > u(x_2 + \Delta x) - u(x_2).$$

Se gli eventi futuri sono incerti e quindi rischiosi, invece, l'utilità da calcolare è un'utilità attesa, definita come la media di tutte le utilità derivanti dai possibili rendimenti futuri, ponderate per le rispettive probabilità di realizzazione:

$$\mathbb{E}[u(X)] = \sum_{i=1}^N u(x_i)p_i.$$

L'utilità attesa è quindi di fatto una media ponderata delle utilità dei possibili esiti dell'investimento e nella formula appena presentata si suppone che la variabile casuale in questione sia discreta.

La funzione dell'utilità attesa viene utilizzata per dare un ordinamento alle alternative rischiose: secondo questo criterio un individuo razionale effettua le scelte in modo tale da massimizzare sempre l'utilità attesa.

Ai fini della costruzione di una funzione di utilità attesa che sia coerente con il comportamento decisionale degli investitori razionali devono essere soddisfatte le proprietà che formuleremo qui di seguito:

- *riflessività*: ogni scelta è almeno tanto preferibile quanto sé stessa, ossia $x_i \succcurlyeq x_i$ ²⁸;
- *transitività*: se $x_h \succcurlyeq x_i$ e $x_i \succcurlyeq x_j$ allora necessariamente $x_h \succcurlyeq x_j$;
- *completezza*: si richiede che l'individuo sia sempre in grado di scegliere, cioè che tra due alternative x_i e x_j sappia dare un giudizio di preferenza o indifferenza tra le stesse, così che risulti $x_i \succcurlyeq x_j$ oppure $x_j \succcurlyeq x_i$.

Tali condizioni devono essere soddisfatte anche nella teoria del consumatore, perché vi sia l'esistenza di una funzione di utilità idonea a rappresentare le

²⁸ La relazione binaria \succcurlyeq indica che x_i è debolmente preferito a x_i .

preferenze dell'individuo. La teoria dell'investimento, invece, esige due ulteriori proprietà:

- *archimedea*: se $x_h \succcurlyeq x_i \succcurlyeq x_j$ allora esistono delle probabilità α e β ²⁹ per le quali si ha $x_h \alpha x_j \succcurlyeq x_i \succcurlyeq x_j \beta x_h$;
- *sostituzione*: se $x_i \succcurlyeq x_j$ allora $x_i \alpha x_h \succcurlyeq x_j \alpha x_h$, per ogni x_h e per ogni probabilità α .

Si può ora affermare che, se una relazione binaria di preferenza gode delle cinque proprietà sopraelencate ed è supposta la razionalità degli investitori, allora esiste una funzione di utilità coerente con la relazione stessa; secondo tale funzione, chiamata $u(X)$, l'alternativa x_1 è preferita a x_2 (entrambe sono variabili casuali discrete) secondo il principio dell'utilità attesa se e solo se l'utilità attesa di x_1 è superiore all'utilità attesa di x_2 . Inoltre la funzione in questione è unica a meno di una trasformazione lineare crescente.

La teoria dell'utilità è stata però oggetto di numerose critiche; molti economisti infatti non condividono il principio di razionalità degli investitori e sostengono che gli individui non sempre adottano comportamenti in linea con le proprietà definite poco sopra.

In letteratura sono state proposte molte funzioni di utilità, come quella esponenziale o logaritmica, ma la Modern Portfolio Theory è compatibile solo con l'utilità quadratica:

$$u(R_p) = R_p - \frac{a}{2} R_p^2, \quad \text{con } a > 0.$$

²⁹ $0 \leq \alpha, \beta \leq 1$.

Il primo addendo è quello legato alla redditività mentre il secondo addendo, con segno negativo, contiene la valutazione soggettiva della rischiosità di un certo investimento. Il parametro a , infatti, indica la personale avversione al rischio dell'individuo. Con riferimento a tale funzione bisogna poi evidenziare due importanti aspetti:

- la derivata prima è positiva, e quindi la funzione è crescente se $R_p < \frac{1}{a}$ e in questo intervallo c'è quindi coerenza con il principio di non sazietà secondo cui ogni individuo tende a massimizzare i propri profitti;
- la derivata seconda è sempre negativa per costruzione, il che implica la concavità della funzione di utilità e quindi l'avversione al rischio insita negli individui.

In particolare, la funzione in questione risulta essere una parabola che però può essere utilizzata per i nostri fini solo per $R_p < \frac{1}{a}$, poiché l'utilità deve necessariamente crescere all'aumentare della ricchezza posseduta. Si veda a tal proposito la Figura 1.5.

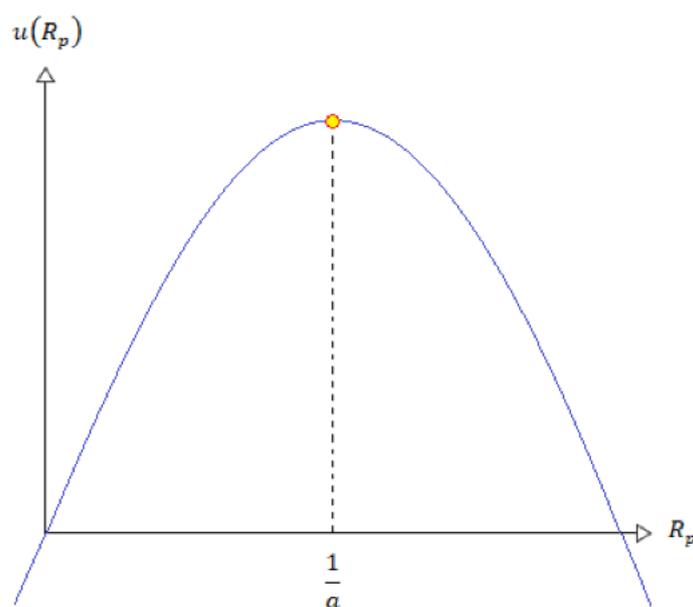


Figura 1.5 - Funzione di utilità quadratica

Calcolando il valore atteso di questa funzione di utilità, si può notare come dipenda solamente dalla media e dalla varianza:

$$\mathbb{E}[u(R_p)] = \mathbb{E}\left(R_p - \frac{a}{2}R_p^2\right) = \mathbb{E}(R_p) - \frac{1}{2}a\mathbb{E}(R_p^2) = \mu - \frac{1}{2}a(\mu^2 + \sigma^2),$$

dove con μ e σ^2 si intendono rispettivamente il valore atteso e la varianza del rendimento del portafoglio. Si noti che, poiché $\sigma^2 = \mathbb{E}(R_p^2) - [\mathbb{E}(R_p)]^2$, nell'espressione sopra risulta $\mathbb{E}(R_p^2) = \mu^2 + \sigma^2$.

La coerenza tra il criterio media-varianza e la funzione di utilità attesa è garantita solo in presenza di una di queste due condizioni mutualmente esclusive:

- quando la funzione di utilità attesa è quadratica;
- quando la distribuzione di probabilità congiunta dei rendimenti è ellittica³⁰.

La seconda condizione è difficile da accettare, in quanto presuppone nella pratica che tutti i rendimenti si distribuiscono secondo una data legge di probabilità, ad esempio Normale, il che risulta spesso in contrasto con l'evidenza empirica. Per questa ragione si preferisce accettare la prima ipotesi, secondo cui la funzione di utilità di tutti gli investitori è quadratica e il suo valore atteso dipende esclusivamente dai parametri media, varianza e avversione al rischio.

³⁰ Sono distribuzioni ellittiche ad esempio la Normale o la t di Student.

La generica curva d'indifferenza³¹ della funzione di utilità quadratica può essere derivata dall'equazione dell'utilità attesa:

$$k = \mu - \frac{1}{2}a(\mu^2 + \sigma^2),$$

dove k è un prefissato valore di utilità attesa e il suo variare porta alla costruzione di varie curve di livello.

Riformulando l'espressione, si può osservare che la generica curva d'indifferenza è una circonferenza di centro $(0, \frac{1}{a})$ e raggio $\sqrt{\frac{1-2ak}{a^2}}$:

$$\sigma^2 + \mu^2 - \frac{2}{a}\mu + \frac{2}{a}k = 0.$$

Ovviamente deve sempre valere il vincolo $\mu < \frac{1}{a}$, altrimenti non è più possibile usare la funzione di utilità quadratica. Sulla medesima curva d'indifferenza sono presenti infinite scelte d'investimento caratterizzate dalla stessa utilità attesa; questo implica che, se un'opzione d'investimento è più rischiosa di un'altra per un investitore, allora, per mantenere invariata l'utilità, la maggiore rischiosità dev'essere accompagnata da un proporzionale aumento del rendimento. Le curve d'indifferenza, quindi, risultano essere crescenti nel piano media-deviazione standard. Esse sono anche convesse in quanto, più il rischio è grande, più l'investitore richiede un rendimento incrementale maggiore per sopportare un aumento della rischiosità; ciò è coerente con l'ipotesi di avversione al rischio degli investitori. Nella Figura 1.6 si possono osservare alcune curve d'indifferenza relative a quattro differenti valori di utilità attesa.

³¹ Le curve d'indifferenza (o di livello) sono l'insieme dei punti aventi stessa utilità attesa. In altre parole, date due scelte d'investimento giacenti sulla stessa curva, esse presenteranno diversi profili di rischio-rendimento ma avranno la stessa utilità per l'investitore.

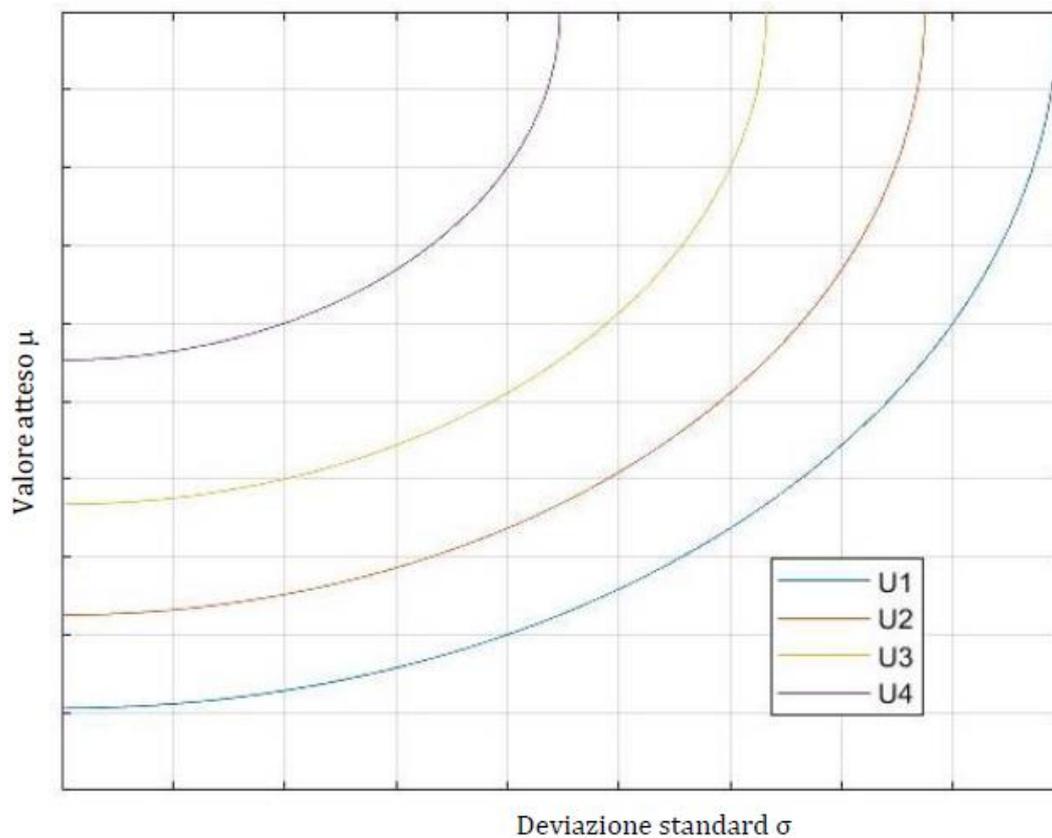


Figura 1.6 - Curve d'indifferenza

Concludendo la presentazione del processo di selezione di portafoglio proposto da Markowitz, ci interessa ora individuare analiticamente il portafoglio ottimo per il singolo investitore. Ricordando che è possibile dimostrare che esiste coerenza tra il criterio media-varianza e la teoria dell'utilità attesa, risulta utile collocare la frontiera efficiente sullo stesso piano delle curve d'indifferenza. In generale, il portafoglio ottimo è quello corrispondente al punto di tangenza tra la frontiera e le curve di livello e quindi è il portafoglio efficiente che allo stesso tempo massimizza l'utilità attesa.

1.3 Critiche alla Modern Portfolio Theory

Negli anni Cinquanta l'intuizione di Markowitz ebbe immediato successo nel mondo finanziario in quanto rivoluzionò l'intero processo di selezione di portafoglio. Il suo contributo fu determinante per la nascita della finanza moderna.

Non mancano tuttavia le critiche che negli anni sono state mosse a questa teoria pionieristica. Il modello proposto da Markowitz, infatti, presenta una serie di limiti che possono distorcere notevolmente i risultati e che riguardano prevalentemente le ipotesi che sono state poste alla base del modello stesso. Le ipotesi sono state ampiamente criticate, in quanto non rispecchiano l'evidenza empirica e sono state quindi giudicate irrealistiche. Di seguito sono presentate alcune delle principali critiche al modello di selezione di Markowitz:

- l'ipotesi di normalità, implicita nel modello, non è generalmente accettata nell'odierno mercato dei capitali. A riprova di ciò è cruciale ricordare uno dei fatti stilizzati osservati da Cont (2001)³² in merito alle distribuzioni dei rendimenti osservate nel mercato. I rendimenti finanziari, infatti tendono ad essere asimmetrici presentando generalmente una skewness³³ negativa: in questo modo la coda sinistra della distribuzione di probabilità risulta essere allungata e le perdite estreme risultano essere molto più elevate dei profitti che si trovano all'estremità opposta della distribuzione. I rendimenti finanziari osservati sul mercato sono inoltre tendenzialmente leptocurtici, ossia

³² Cont, R. (2001), *Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues*, *Quantitative Finance*, 1, 223-236.

³³ La skewness è un indice di asimmetria di una distribuzione di probabilità, che misura l'asimmetria dei valori di una variabile casuale attorno alla propria media. In caso di perfetta simmetria, come nella distribuzione Normale, la skewness è pari a 0.

presentano una curtosi³⁴ elevata che implica una distribuzione più appuntita rispetto a quella Normale ma con delle code più spesse, chiamate *fat tails*;

- la varianza è stata ampiamente criticata come misura di rischio perché tratta i rendimenti positivi alla stessa stregua dei rendimenti negativi, quando invece gli investitori sono soliti pesare maggiormente il lato della distribuzione relativo alle perdite, poiché sono avversi al rischio. È corretto usare la varianza come misura di rischio solo se i rendimenti presentano una distribuzione simmetrica, cosa che nella realtà non si verifica spesso;
- le assunzioni di base per cui non esistono né tasse né costi di transazione e tutti i titoli sono perfettamente divisibili sono alquanto irrealistiche, poiché nel mercato vi sono numerosi attriti e costi che devono essere tenuti in considerazione del processo di selezione del portafoglio. I titoli inoltre non sono perfettamente divisibili, in quanto vi sono dei lotti minimi di transazione di cui bisogna tenere conto.

Un altro limite importante fu messo in evidenza dall'economista William Sharpe³⁵, partendo dal presupposto che il numero di input richiesto da questo modello di selezione di portafoglio è molto elevato. La diversificazione è il principio cardine di questa teoria e induce gli investitori a costruire portafogli con un notevole numero di titoli. Aggiungendo titoli al portafoglio, però, aumenta esponenzialmente il numero di input richiesto. Questo modello, infatti, necessita della stima di N medie dei rendimenti, N varianze dei

³⁴ L'indice di curtosi è uno degli indici relativi alla forma di una distribuzione ed è una misura dello spessore delle code di una funzione di densità, ovvero indica il grado di appiattimento di una distribuzione. La distribuzione Normale presenta un valore di tale indice pari a 3.

³⁵ William F. Sharpe, *A Simplified Model for Portfolio Selection*, Management Science, Vol. 9, No. 2, University of Washington, January, 1963, pp. 277 – 293.

rendimenti e $\frac{N(N-1)}{2}$ covarianze tra i rendimenti. Per cui, a titolo esemplificativo, un portafoglio composto da soli 200 titoli richiede la stima di 20.300 parametri. Sharpe osservò che la stima di un numero così elevato di parametri presenta dei seri problemi per due ordini di motivi:

- il primo, molto più evidente negli anni Cinquanta del secolo scorso, è il tempo di calcolo necessario per effettuare tutte le stime;
- il secondo deriva dal fatto che, poiché bisogna effettuare un gran numero di stime puntuali e queste sono affette da un certo errore di stima, allora tali errori si combinano e si amplificano impattando sulla composizione del portafoglio selezionato.

Con l'obiettivo di risolvere questi due problemi, Sharpe sviluppò il *modello diagonale*, un modello per l'analisi di portafoglio che richiede la stima di un numero di parametri inferiore; in questo modo si può beneficiare di un risparmio in termini di tempo computazionale e di numero di errori di stima.

Il modello di Markowitz, inoltre, parte dall'assunzione che tutti gli investitori siano razionali, e in quanto tali soddisfino sia l'assioma di non sazietà sia quello di avversione al rischio. Secondo tale teoria l'investitore è sempre alla ricerca del punto di ottimo, massimizzando il rendimento del proprio investimento per un dato livello di rischio e sfruttando al meglio tutte le informazioni disponibili. Di fatto però molte ricerche hanno dimostrato che gli individui sono afflitti da non piena razionalità e da incoerenza e incompetenza nel momento in cui si confrontano con eventi caratterizzati da incertezza. È la finanza comportamentale che si è proposta di spiegare in che modo emozioni ed errori cognitivi possano influenzare gli investitori e i loro processi decisionali. Per

costruire dei modelli coerenti con l'attuale realtà finanziaria è quindi necessario tenere conto anche dell'aspetto psicologico e comportamentale degli individui, complicando così notevolmente la struttura matematica dei modelli stessi.

In conclusione di questo capitolo si riportano tre ulteriori problematiche relative al processo di selezione di portafoglio proposto da Markowitz:

- *scelta del valore di π* - ossia del rendimento desiderato - necessario per la procedura di ottimizzazione: questo è un aspetto cruciale in quanto per l'investitore medio, che non ha un solido background economico-finanziario, risulta difficile scegliere un target return adeguato alle proprie esigenze. C'è quindi un problema di interpretabilità dei modelli e dei risultati da parte dell'investitore medio. Tra le molte soluzioni possibili quest'ultimo potrebbe ad esempio scegliere di investire nel portafoglio che dà un rendimento maggiore tra il global minimum variance portfolio e l'equally-weighted. Con questo criterio di scelta la procedura decisionale dell'investitore risulta di gran lunga più agevole rispetto al dover scegliere autonomamente un target return adeguato;
- *rischio di stima relativo ai parametri del modello*: dai test empirici risulta che il portafoglio è più suscettibile agli errori di stima delle medie che delle varianze e covarianze. Per questa ragione investire nel global minimum variance portfolio o nell'equally-weighted risulta sensato in quanto questi sono i portafogli con il minor numero di stime e quindi di errori al loro interno. Il rischio di stima ha una rilevanza fondamentale in quanto piccoli errori di stima relativi al singolo parametro di un singolo titolo si amplificano notevolmente a livello di portafoglio, distorcendo sensibilmente la selezione;
- *instabilità delle soluzioni*: i mercati finanziari non sono stazionari, perciò i processi che generano i dati non sono costanti nel tempo e quindi anche i

portafogli, essendo strettamente dipendenti dalla realtà finanziaria, possono perdere la loro efficienza nel tempo. È quindi necessaria una procedura di revisione di portafoglio idonea a cogliere i cambiamenti rilevanti che avvengono nei mercati finanziari. Le percentuali investite nei singoli titoli che compongono il portafoglio devono quindi essere periodicamente aggiustate, tenendo in considerazione le dinamiche correnti di mercato.

Illustrati i limiti della Modern Portfolio Theory, possiamo ora affermare che l'approccio di Markowitz, seppur sia un pilastro indissolubile della finanza moderna, presenta non pochi problemi applicativi nella realtà finanziaria odierna, poiché le assunzioni alla base del modello non corrispondono a quello che effettivamente viene osservato nei mercati. Per queste ragioni sono necessarie delle modifiche al modello di portafoglio che tengano in considerazione tra l'altro anche il fattore psicologico e non razionale insito negli individui e che il modello di Markowitz non riesce a cogliere.

CAPITOLO 2

DALLE MISURE DI RISCHIO CLASSICHE ALLE MISURE DI RISCHIO COERENTI

2.1 I limiti delle misure di rischio classiche

Il *rischio* è un concetto cruciale nella teoria finanziaria e nel corso degli anni ha subito varie modifiche e adattamenti, a causa dei mutamenti dei mercati finanziari e dell'evoluzione delle teorie volte a studiare i fenomeni economici.

La teoria di portafoglio proposta da Markowitz¹ si basa sul concetto di varianza², che è una misura di rischio classica. La sua radice quadrata è la deviazione standard (o scarto quadratico medio), nota anche come volatilità, ed è una misura di dispersione dei rendimenti intorno al valor medio. Rileva pertanto l'incertezza che regna sul mercato. Non è detto però che ad elevata volatilità corrisponda un'elevata probabilità di perdita, in quanto vengono tenuti in considerazione sia gli scostamenti negativi che quelli positivi (i profitti).

Volendo misurare l'effettiva rischiosità di un portafoglio risulta quindi più utile focalizzarsi sulle potenziali perdite ossia sul lato sinistro della distribuzione di probabilità dei rendimenti, poiché i profitti sono finanziariamente auspicabili in un investimento. In questo modo si giunge al concetto di semi-varianza e vengono considerati solo gli scostamenti negativi dei rendimenti da un certo

¹ H. Markowitz, *Portfolio selection: efficient diversification of investments*, New York, John Wiley & Sons, 1959.

² Si ricordi che, nel caso di variabili casuali discrete, la varianza (indicata anche con σ^2) è pari a: $Var(X) = \sum_{i=1}^N [x_i - E(X)]^2 p_i$, dove x_i è la i -esima realizzazione della variabile casuale X con la rispettiva probabilità p_i .

valore di riferimento (che può essere lo 0 oppure il valore medio). Nel caso di variabili casuali discrete la semi-varianza viene definita nel seguente modo:

$$\text{semi} - \text{Var}(X) = \sum_{i=1}^N (\min\{0, x_i - \mathbb{E}(X)\})^2 p_i.$$

Si riesce così a misurare il cosiddetto *downside risk*, ossia la parte negativa della volatilità dell'investimento.

Un'altra misura di rischio classica è la mean absolute deviation che, nel caso di variabili casuali discrete, viene definita dalla seguente formula:

$$\text{MAD}(X) = \sum_{i=1}^N |x_i - \mathbb{E}(X)| p_i.$$

A differenza della varianza che considera la media dei quadrati degli scostamenti dal valor medio, la mean absolute deviation considera la media dei valori assoluti di tali scostamenti.

Le misure appena descritte sono più che altro indicatori di variabilità piuttosto che di rischio, in quanto ciascuna di esse presenta notevoli problemi applicativi come misura di rischio, specie nei mercati finanziari odierni.

Per quanto riguarda la varianza è anche opportuno ricordare quanto emerso dal matematico Szegö (2002): tale misura di dispersione può essere usata efficacemente come misura di rischio se le distribuzioni dei rendimenti sono simmetriche. L'evidenza empirica ci suggerisce invece una marcata asimmetria dei rendimenti³, il che generalmente tende ad essere sfavorevole per l'investitore in quanto la massa di probabilità tende a concentrarsi

³ I primi lavori empirici sono stati effettuati da Mandelbrot (1963) e Fama (1965) e hanno escluso l'ipotesi di normalità distributiva nelle serie storiche dei rendimenti finanziari.

maggiormente sul lato delle perdite piuttosto che su quello dei profitti. Minimizzare la varianza in presenza di asimmetria positiva significa ridurre notevolmente le possibilità di guadagno. Di conseguenza tale misura, seppur presenti notevoli vantaggi analitici, comporta diversi problemi applicativi di non poco conto.

La semi-varianza invece risulta utile quando il rischio viene percepito non tanto come dispersione intorno ad un certo valore ma viene legato alla possibilità di subire delle perdite.

Infine la mean absolute deviation presenta caratteristiche e problematiche molto simili alla varianza e sono entrambe sensibili agli *outliers*⁴. Quest'ultima misura, per di più, non è agevolmente trattabile per la presenza dell'operatore valore assoluto e di conseguenza, come la semi-varianza, comporta problemi applicativi dati dalla presenza di un punto di non derivabilità.

Nella nostra disamina delle misure di rischio classiche possiamo perciò concludere che esse sono di per sé degli indici di dispersione e risultano generalmente non idonee come misure di rischio per un problema di ottimizzazione di portafoglio, tenendo conto della discrepanza tra la teoria finanziaria che si è andata a formare nella seconda metà del secolo scorso e l'effettivo andamento dei mercati odierni.

2.1.1 *Proprietà delle misure di rischio*

Nel corso degli anni si è sempre più sentita la necessità di definire univocamente i requisiti che deve possedere un'opportuna misura di rischio finanziario, concretamente applicabile e coerente con le attuali dinamiche di mercato.

⁴ In statistica vengono definiti così dei valori anomali e quindi distanti dalle altre osservazioni disponibili.

Seppure sia possibile specificare a priori le caratteristiche che una buona misura di rischio dovrebbe possedere, il concetto di rischio è alquanto soggettivo⁵ e di conseguenza risulta utopico trovare una misura che possa risolvere il problema, comune ad ogni investitore, di massimizzazione dell'utilità attesa. Tuttavia, restringendo il campo degli investitori a quello degli agenti razionali, è possibile determinare delle caratteristiche che una misura di rischio è auspicabile possieda.

Tra i vari tentativi di dare una definizione chiara e precisa di misura di rischio si ricorda quanto proposto dai matematici Szegö (2002), Ortobelli *et al.* (2005) e Rachev *et al.* (2008): una misura di rischio è una funzione ρ che assegna un valore numerico non negativo ad una variabile casuale X interpretabile come il rendimento futuro, $\rho: X \rightarrow \mathbb{R}$.

Possiamo ora individuare le proprietà di base che una misura di rischio deve soddisfare affinché una funzione ρ sia definibile come tale, seppur non siano le uniche a dover essere tenute in considerazione: a breve verrà illustrato il concetto di coerenza di una misura di rischio, così come è stato definito nel lavoro di Artzner *et al.* (1999).

Partendo da quelle più elementari ed intuitive, le proprietà di cui si è appena accennato sono le seguenti:

- *positività*: una misura di rischio associa ad una variabile casuale un valore strettamente positivo, al più nullo nel caso in cui non vi sia alcuna forma di casualità. Valori negativi non hanno senso e sarebbero in contraddizione con la definizione di rischio citata poco sopra;
- *linearità*: soprattutto nel caso in cui si vogliono risolvere problemi di ottimizzazione di grandi dimensioni, la complessità computazionale del

⁵ Il rischio è legato alla percezione dell'incertezza da parte del singolo investitore, nonostante alcuni autori distinguano tra rischio (quantità misurabile) e incertezza (quantità non misurabile).

problema di programmazione matematica può essere attenuata se la misura di rischio e i vincoli sono funzioni lineari. Il successo di alcune misure di rischio è legato alla trattabilità computazionale che deriva dal problema di ottimizzazione “linearizzato”, in quanto è più semplice risolvere un problema di programmazione lineare piuttosto che uno di programmazione quadratica o di grado superiore;

- *convessità*: una misura di rischio è convessa se, dati due rendimenti aleatori R_X e R_Y ed un parametro $\vartheta \in [0; 1]$, vale la seguente relazione:

$$\rho(\vartheta R_X + (1 - \vartheta)R_Y) \leq \vartheta\rho(R_X) + (1 - \vartheta)\rho(R_Y).$$

È una proprietà che sottolinea l'importanza della diversificazione, il processo mediante il quale il rischio complessivo di un portafoglio tende a diminuire asintoticamente all'aumentare del numero dei titoli e tutt'al più non aumenta. La presenza in un portafoglio di strumenti finanziari diversi tra loro riduce sistematicamente il rischio complessivo in quanto l'eventuale correlazione positiva che esiste tra i diversi titoli non è mai perfetta. Un portafoglio adeguatamente diversificato sopporta solamente il rischio di mercato, ossia quello relativo alla variazione sfavorevole dell'insieme dei prezzi dei titoli trattati in un dato mercato, che per sua natura è ineliminabile. La diversificazione elimina invece il rischio specifico ossia quello tipico dei singoli titoli, le variazioni dei rendimenti dei quali tendono a compensarsi tra loro. La proprietà di convessità è inoltre molto importante nei processi di ottimizzazione poiché garantisce che un punto di minimo locale sia anche un punto di minimo globale (Uryasev, 2000). Essa è inoltre soddisfatta anche indirettamente nel momento in cui la misura di rischio in questione soddisfa le due seguenti ed ulteriori proprietà, che verranno adeguatamente riprese e spiegate quando si affronterà il tema degli assiomi di coerenza:

- *subadditività*⁶: secondo questa proprietà la rischiosità complessiva di un portafoglio non dovrebbe mai superare la somma dei singoli indici di rischiosità associati ai vari titoli che lo compongono. La subadditività è perciò strettamente legata al concetto di diversificazione e può essere formalizzata nella maniera seguente:

$$\rho(R_X + R_Y) \leq \rho(R_X) + \rho(R_Y);$$

- *positiva omogeneità*⁷: la seguente proprietà garantisce un legame direttamente proporzionale tra l'ammontare investito in un titolo a rendimento aleatorio e la rischiosità associata a quell'investimento. Può essere formalmente definita nella maniera seguente:

$$\rho(\alpha R_X) = \alpha \rho(R_X) \quad \forall \alpha \in \mathbb{R}^+.$$

2.2 Le misure di rischio premoderne

Preso la consapevolezza che la varianza e le altre misure classiche non sono delle vere e proprie misure di rischio, negli anni si è sentita sempre di più la necessità di una certa innovazione finanziaria in quest'ambito. La ricerca di nuove misure di rischio si è imbattuta inizialmente in alcune misure che, anche se non particolarmente idonee, hanno ricoperto e ricoprono un ruolo estremamente rilevante. Ricordiamo, ad esempio, il *Value-at-Risk* (d'ora in avanti *VaR*), che appartiene al gruppo delle misure di rischio premoderne.

⁶ Verrà ripresa e illustrata più accuratamente nelle seguenti pagine, descrivendo gli assiomi di coerenza.

⁷ Anche questa proprietà verrà ripresa descrivendo gli assiomi di coerenza.

2.2.1 Vantaggi e limiti del Value-at-Risk

La nozione di VaR è piuttosto semplice e di facile intuizione. La definizione è la seguente: il VaR ad un livello di confidenza $\alpha \in [0; 1]$ ⁸, fissato un certo intervallo temporale (noto come *holding period*), indica qual è la massima perdita potenziale in cui il portafoglio sottostante può incorrere nell' $\alpha\%$ dei casi ed in un arco temporale pari all'*holding period*⁹.

Il VaR è di per sé nato come misura di rischio di mercato¹⁰ associato ad un'attività finanziaria: con un dato livello di confidenza, infatti, esso rappresenta la massima perdita ipotizzabile nell'*holding period* derivante dall'attività finanziaria in questione. L'orizzonte temporale riflette il tempo minimo necessario a smobilizzare l'investimento in caso di perdita, che può variare a seconda della liquidità del mercato in cui si sta operando.

Tale misura di rischio riveste inoltre un ruolo fondamentale all'interno dell'industria del credito e del risparmio gestito, ai fini della determinazione dei requisiti patrimoniali previsti dalle normative di vigilanza. Ad esempio, è in base ai rischi di mercato cui una banca è esposta, che sono misurati col VaR, che le autorità di vigilanza richiedono di detenere un ammontare minimo di capitale per far fronte a suddetti rischi. Se una banca non dovesse disporre di modelli interni¹¹ per il calcolo del VaR, tali modelli vengono suggeriti

⁸ Nella maggior parte dei casi è fissato a 0.95 o 0.99.

⁹ Solitamente pari ad 1 giorno o a 10 giorni.

¹⁰ È il rischio relativo agli effetti imprevisi sul valore di mercato di attività e passività prodotti da variazioni dei tassi di interesse, dei tassi di cambio e da altri prezzi delle attività.

¹¹ Nel 2000 la Banca d'Italia, seguendo quanto approvato nell'emendamento all'Accordo di Basilea, ha emanato le direttive di vigilanza che consentono alle banche di calcolare i requisiti patrimoniali a fronte dei rischi di mercato utilizzando i propri modelli interni, che devono però rispettare determinati requisiti qualitativi e quantitativi. Si è scelto il VaR come metodo di misurazione del rischio. L'emendamento ha previsto inoltre la possibilità di utilizzare una metodologia di

direttamente dalla rispettiva normativa, che nel settore bancario è Basilea III¹² e predispone dei processi di calcolo rigidamente strutturati e standardizzati.

Il VaR è stato concepito inizialmente dalla banca d'investimento americana J.P. Morgan¹³ e presenta il grande vantaggio di esprimere il rischio nella stessa unità di misura del capitale investito e quindi in termini monetari, a prescindere dallo strumento sottostante.

Questa misura di rischio non è però esente da limiti in quanto se le perdite sottostanti non seguono distribuzioni gaussiane, allora il VaR diventa uno strumento instabile.

Inoltre non si riesce a stimare l'ammontare delle perdite superiori rispetto alla soglia indicata da tale misura. In altre parole il VaR non tiene in considerazione ciò che avviene oltre la soglia nella coda sinistra della distribuzione di probabilità: questa misura di rischio può dunque far giungere alla fuorviante conclusione che due portafogli abbiano la stessa rischiosità, mentre di fatto la distribuzione dei rendimenti di uno dei due è caratterizzata dalla presenza di eventi di coda che, seppur poco probabili, comportano perdite nettamente più elevate rispetto a quanto si possa verificare probabilisticamente nell'altro portafoglio. Il VaR ha quindi la tendenza ad essere troppo ottimista nelle sue stime mentre invece è sempre consigliato adottare un approccio più prudente, per proteggersi da eventuali perdite maggiori rispetto a quanto era stato

calcolo standardizzata per le banche non dotate delle competenze necessarie ad implementare modelli basati sul VaR.

¹² Gli Accordi di Basilea sono linee guida in materia di requisiti patrimoniali delle banche, redatte dal Comitato di Basilea, costituito dagli enti regolatori del G10 (composto attualmente da undici paesi), oltre che dal Lussemburgo, allo scopo di perseguire la stabilità monetaria e finanziaria. Le nuove regole introdotte con Basilea III definiscono nuovi standard internazionali per l'adeguatezza patrimoniale delle banche e nuovi vincoli di liquidità.

¹³ Si veda *RiskMetrics™ - Technical Document*, J.P. Morgan/Reuters, Fourth Edition (1996). È una metodologia analitica (parametrica) introdotta nel 1994 da J.P. Morgan con lo scopo di calcolare il VaR di un portafoglio di investimenti e confrontare così il rischio associato a posizioni diverse.

preventivato. Nelle pagine seguenti introdurremo delle misure di rischio che considerano le perdite oltre il VaR, distinguendo così quelle poco superiori da quelle che potenzialmente possono azzerare tutti i profitti conseguiti.

Infine è estremamente difficile, se non impossibile, stimare le distribuzioni di probabilità delle posizioni sottostanti e da ciò ne consegue che è molto difficile ottenere una misura precisa del VaR di un portafoglio.

Nella Figura 2.1 viene riportata l'esemplificazione grafica di quanto detto, necessaria ai fini della comprensione del Value-at-Risk, misura di rischio di per sé non concettualmente complicata ma cruciale in quanto getta le basi per tutta una serie di misure e di concetti che verranno illustrati successivamente.

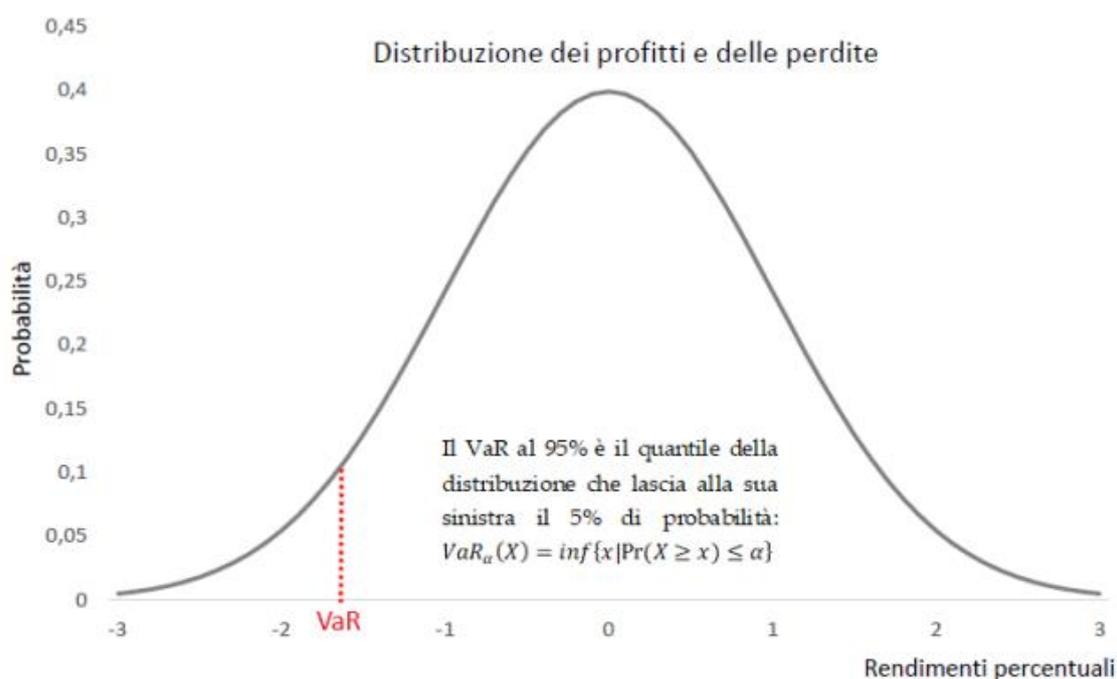


Figura 2.1 - Il Value-at-Risk come $(1-\alpha)$ -esimo percentile della distribuzione dei rendimenti

Dal punto di vista probabilistico, il VaR - considerato un livello di confidenza α - è quel valore che soddisfa la seguente uguaglianza: $\Pr(L > VaR_{\alpha}) = 1 - \alpha$, dove L è una generica variabile casuale che indica le perdite.

Si comprende ora bene uno dei limiti del VaR, già accennato precedentemente. Esso è una misura soglia e in quanto tale esprime l'entità della massima perdita potenziale, dato un certo livello di confidenza statistica. Non fornisce però alcuna indicazione sulla coda sinistra della distribuzione che eccede tale soglia e quindi rimane un vuoto informativo sull'ammontare delle perdite che si verificherebbero in caso questa soglia venisse superata.

Per questo motivo il VaR non è una misura idonea all'attività di *risk management*: tale attività infatti predilige un atteggiamento prudentiale ai fini della misurazione del rischio associato ad un certo portafoglio e ciò porta a sovrastimare l'ammontare di perdite che si potranno verificare in futuro. Il VaR invece è una misura che tende a sottostimare le perdite, contrapponendosi quindi ai principi base della gestione del rischio.

Due portafogli con lo stesso VaR possono dunque subire perdite di entità nettamente differente nel peggior $(1-\alpha)\%$ dei casi. Se le perdite si concentrano significativamente nelle code della distribuzione di probabilità, allora si parla di *tail effect*. Possiamo rappresentare tre possibili scenari, rispettivamente in Figura 2.2, Figura 2.3, Figura 2.4, dove il *tail effect* non si manifesta, si manifesta in maniera moderata oppure estrema. Si noti che, supponendo che il VaR con una certa confidenza sia lo stesso, la perdita attesa condizionata al superamento del VaR è nettamente superiore nel terzo scenario.

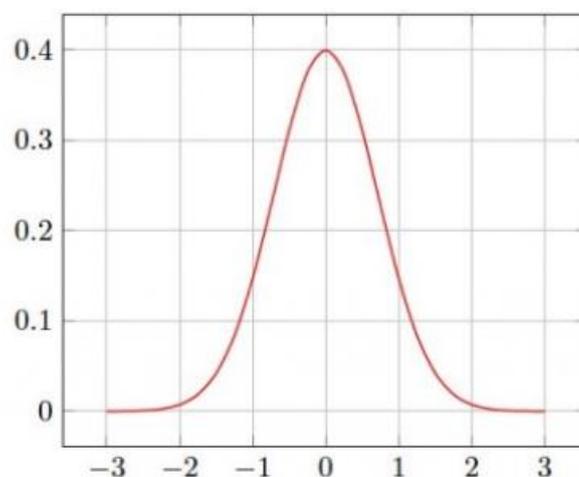


Figura 2.2 - Primo scenario: assenza di tail effect

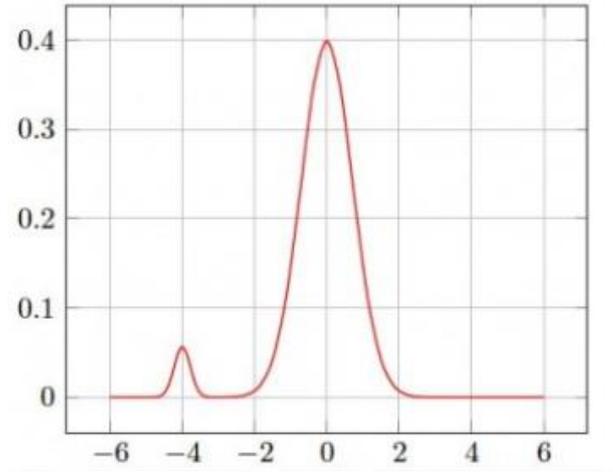


Figura 2.3 - Secondo scenario: tail effect moderato

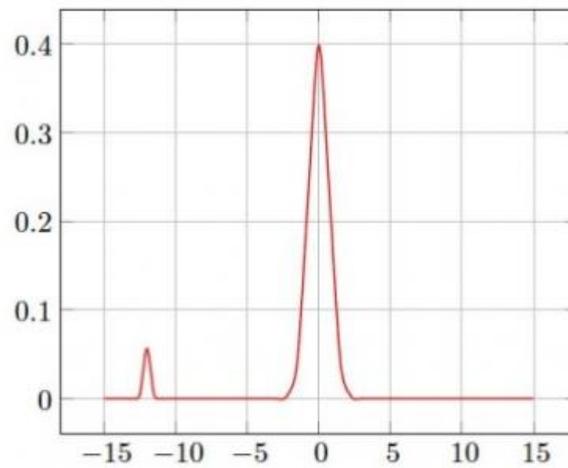


Figura 2.4 - Terzo scenario: tail effect estremo

Infine è necessario evidenziare un'ulteriore lacuna presente in tale misura di rischio: il VaR non soddisfa l'assioma di subadditività e quindi non può essere considerato una misura di rischio coerente¹⁴. Applicando il VaR ad un portafoglio finanziario si potrebbe così verificare una situazione paradossale in cui il rischio complessivo di tale portafoglio superi la somma delle singole fonti di rischio provenienti da ciascun strumento finanziario. Risulta evidente perciò che, come sostenuto da Artzner *et al.* (1999), il VaR non incoraggia e addirittura

¹⁴ Si vedano le seguenti pagine per la definizione di *misura di rischio coerente*.

talvolta proibisce la diversificazione, che invece dovrebbe essere promossa in quanto fondamentale ai fini della gestione del rischio di portafoglio.

2.3 Le misure di rischio coerenti

A questo punto, compresi i vantaggi e i limiti del Value-at-Risk, risulta necessario proseguire l'analisi delle proprietà desiderabili che una misura di rischio deve possedere affinché possa essere usata efficacemente ed efficientemente in un problema di ottimizzazione di portafoglio. Affrontiamo così il tema dei cosiddetti assiomi di coerenza, individuati nel lavoro di Artzner *et al.* (1999). Gli autori non hanno voluto definire una specifica misura di rischio, ma piuttosto hanno individuato delle proprietà che caratterizzano un'ampia classe di misure di rischio, definite *coerenti*. La comprensione e l'utilizzo di queste ultime è a dir poco cruciale ai fini della finanza quantitativa moderna, in quanto forniscono un supporto decisionale decisamente appropriato ai fini di una corretta ottimizzazione di portafoglio.

2.3.1 Gli assiomi di coerenza

Una misura di rischio coerente è una misura che soddisfa i quattro assiomi di coerenza riportati di seguito. Premettiamo innanzitutto che tali assiomi sono maggiormente intuitivi se si ragiona in termini monetari invece che percentuali. La violazione di anche solo uno di questi assiomi porta a risultati erronei e paradossali: per questo Acerbi e Tasche (2002) considerano essere misure di rischio finanziario vere e proprie solo quelle coerenti. Gli assiomi di coerenza definiscono il concetto di rischio a livello nettamente più operativo rispetto a quanto mai era stato fatto prima e sono i seguenti:

1. *Invarianza transazionale*: assicura che investendo una parte del capitale a disposizione in un titolo privo di rischio (a rendimento certo), il rischio complessivo del portafoglio si contrae proporzionalmente alla quota α allocata nello strumento risk-free:

$$\rho(X + \alpha) = \rho(X) - \alpha \quad \forall v. c. X, \alpha \in \mathbb{R}.$$

Tale assioma implica che $\rho(X + \rho(X)) = 0$. Aggiungendo quindi una quantità priva di rischio esattamente pari a $\rho(X)$ alla posizione rischiosa X si ottiene un'entità a rischio nullo, il che è coerente con l'interpretazione operativa di ρ quale minima quantità positiva di denaro che l'investitore deve aggiungere alla posizione iniziale al fine di creare uno strumento "accettabile" e quindi privo di rischio.

2. *Subadditività*: questo assioma mette in luce una proprietà fondamentale che le misure di rischio dovrebbero possedere e riguarda l'aggregazione di più fonti di rischio. L'indice di rischiosità di un portafoglio non dovrebbe mai superare la somma dei singoli indici di rischiosità associati ai vari titoli che lo compongono. Gli stessi autori evidenziano come «*a merger does not create extra risk*» (Artzner *et al.*, 1999). La subadditività è perciò strettamente legata al concetto di diversificazione in quanto, come in parte era già stato rilevato precedentemente, la diversificazione contribuisce alla riduzione del rischio complessivo di portafoglio soltanto nel caso in cui il rischio di una singola posizione¹⁵ venga misurato attraverso una misura subadditiva:

$$\rho(X + Y) \leq \rho(X) + \rho(Y) \quad \forall v. c. X, Y.$$

¹⁵ Si parla anche di *profilo di rischio stand-alone*.

3. *Positiva omogeneità*: garantisce un legame direttamente proporzionale tra l'ammontare investito in un titolo a rendimento aleatorio e la rischiosità associata a quell'investimento. Dato un numero non negativo λ , il rischio relativo a λ volte un investimento dev'essere pari a λ volte il rischio dell'investimento stesso. Questo assioma può essere riassunto nella maniera seguente:

$$\rho(\lambda X) = \lambda \rho(X) \quad \forall v. c. X, \lambda \geq 0.$$

4. *Monotonicità*: questo assioma sostiene che se un individuo preferisce l'investimento X a Y per un qualche motivo, allora la misura di rischio deve essere in grado di recepire la preferenza di quell'individuo. Formalmente, date due variabili casuali X e Y con $X \succcurlyeq Y$ (X è preferita a Y), allora il rischio associato alla variabile X sarà minore del rischio associato alla variabile Y :

$$\rho(X) \leq \rho(Y) \quad \forall v. c. X, Y \text{ con } X \succcurlyeq Y.$$

Nel caso limite in cui nessuno dei due investimenti è preferito all'altro e quindi sono perfettamente equivalenti, l'assioma di monotonicità assicura l'uguaglianza tra le rischiosità degli stessi. Qualora esistono opportunità di arbitraggio¹⁶ nei mercati finanziari, una misura di rischio che soddisfa questo assioma aiuta a sfruttarle, ovvero a realizzare un

¹⁶ L'arbitraggio è un'operazione che consente di ottenere un profitto certo a rischio nullo (si parla di *pasto gratis*). Le opportunità di arbitraggio ci sono ogniqualvolta una stessa attività finanziaria è scambiata su due diversi mercati a prezzi diversi tra loro, oppure quando due attività finanziarie perfettamente sostituibili tra loro vengono scambiate sullo stesso mercato a prezzi diversi. Gli investitori che individuano tali opportunità possono agire immediatamente acquistando il titolo venduto a un prezzo inferiore e vendendo quello il cui prezzo è maggiore. In questo modo si sfrutta la differenza di prezzo tra due strumenti finanziari e come effetto si ha il riallineamento dei prezzi che erano temporaneamente disallineati.

profitto senza rischiare nulla. Supponendo ad esempio che lo stesso titolo sia prezzato diversamente in due differenti mercati, l'assioma di monotonicità assicura che una misura di rischio coerente attribuisce al titolo con una quotazione inferiore - quindi meno redditizio - una rischiosità maggiore. Di conseguenza un investitore può comprare il titolo nel mercato dove presenta un rischio minore e contemporaneamente venderlo nel mercato in cui presenta una rischiosità più alta, ottenendo così un profitto certo. In questi termini apparentemente viene contraddetto uno dei principi fondamentali della finanza secondo cui ad un maggior rischio è associato un maggior rendimento: ciò però vale solo nel caso in cui non vi siano possibilità di arbitraggio nei mercati.

Possiamo quindi concludere che, affinché una misura di rischio possa essere definita coerente, deve necessariamente rispettare questi quattro assiomi appena descritti. Prima di passare alla disamina di alcune misure coerenti vogliamo ricordare che, sotto un profilo strettamente analitico, la subadditività è una proprietà essenziale ai fini dell'ottimizzazione di portafoglio. Questo perché essa, assieme alla positiva omogeneità, implica la proprietà di convessità; quest'ultima per giunta porta il problema di ottimizzazione e quindi di minimizzazione del rischio ad avere un'unica soluzione.

Infine, bisogna evidenziare il fatto che il mondo accademico fin da subito si è dimostrato favorevole all'adozione di queste nuove misure di rischio. Non si può dire lo stesso però dei regolatori e *practitioners*¹⁷, che sono ancora abbastanza sordi di fronte a quest'innovazione finanziaria e continuano a perseverare nell'utilizzo di misure di rischio non coerenti come il VaR, con tutte

¹⁷ Si intende coloro che operano nei mercati, ossia i professionisti.

le conseguenze del caso soprattutto quando vengono aggregate diverse fonti di rischio.

2.3.2 TCE e WCE

Artzner *et al.* (1999), all'interno del medesimo lavoro in cui hanno definito gli assiomi di coerenza, hanno anche fornito i primi esempi concreti di misure di rischio che soddisfano, anche se con qualche eccezione, le quattro proprietà sopraelencate.

Gli autori si sono focalizzati su due misure, note come *Tail Conditional Expectation (TCE)* e *Worst Conditional Expectation (WCE)*. Essi hanno dimostrato che è soddisfatta la relazione $TCE_\alpha \leq WCE_\alpha$ ¹⁸ e per comprendere il significato di quanto enunciato si riportano di seguito le definizioni:

- la Tail Conditional Expectation (anche nota come “TailVaR”) è una misura di rischio coerente definita come:

$$TCE_\alpha(X) \stackrel{\text{def}}{=} -\mathbb{E}[X|X \leq -VaR_\alpha(X)]^{19};$$

- la Worst Conditional Expectation è anch'essa una misura coerente definita però come:

$$WCE_\alpha(X) \stackrel{\text{def}}{=} -\inf\{\mathbb{E}[X|A] \mid \Pr[A] > \alpha\}^{20}.$$

¹⁸ Nelle seguenti formule si usa la notazione α per indicare la percentuale (di solito 1% o 5%) relativa ai possibili scenari peggiori per il portafoglio che stiamo analizzando. In questi termini con VaR_α si intende la massima perdita potenziale che il portafoglio può subire nell' $(1 - \alpha)\%$ dei casi.

¹⁹ Essa calcola qual è la perdita media se essa risulta superiore alla soglia corrispondente al VaR. Il segno meno ha il solo scopo di rendere positiva la misura, in quanto normalmente si lavora con le perdite.

Seppur siano state entrambe definite come misure coerenti bisogna fare una distinzione, in quanto la WCE è stata introdotta proprio per supplire a un limite della TCE. Quest'ultima, infatti, come è stato evidenziato da Acerbi e Tasche (2002), è una misura più maneggevole da utilizzare in ambito applicativo ma non soddisfa sempre gli assiomi di coerenza in quanto può non essere subadditiva²¹.

La WCE invece, sebbene risponda a tutti gli assiomi di coerenza, è ampiamente diffusa solo in ambito teorico in quanto richiede la conoscenza dell'intero spazio di probabilità sottostante²².

Dal punto di vista finanziario queste due misure tengono conto di "*how bad is bad*" perché si focalizzano sulla forma della coda sinistra della distribuzione dei rendimenti, ossia si concentrano sulle perdite calcolando il loro valor medio condizionato al fatto che tali perdite risultino superiori ad un certo valore.

A questo punto sorge spontaneo un rapido confronto con il Value-at-Risk, il quale non fornisce alcuna indicazione sulla natura delle perdite oltre un valore soglia che è rappresentato dallo stesso VaR_α . Con l'introduzione degli assiomi di coerenza si è giunti ad un radicale cambiamento del quesito su cui è necessario interrogarsi per comprendere l'entità del rischio a cui si è esposti. In altri termini, quasi tutte le misure definite "coerenti" non guardano alla massima perdita potenziale nell' $\alpha\%$ dei casi, bensì a qual è la perdita attesa nel peggior $(1-\alpha)\%$ dei casi. Tali misure non si concentrano più perciò su un

²⁰ Tale formula calcola il valore medio di tutte le perdite date dall'evento A , la cui probabilità di accadimento sia superiore ad α e tiene in considerazione esclusivamente il valore corrispondente alla perdita maggiore.

²¹ La coerenza della TCE è garantita solo restringendo il campo d'analisi alle funzioni di distribuzione di probabilità continue, mentre potrebbe non essere garantita nel caso più generale. Questo è il motivo che ha spinto Artzner *et al.* (1999) all'introduzione della WCE.

²² L'implementazione della WCE richiede anche la specificazione di un metodo in grado di individuare tutti gli scenari con probabilità maggiore o uguale al livello di significatività α fissato.

determinato valore soglia, inidoneo a rappresentare l'entità di rischio che ci si sta assumendo, ma si concentrano sull'intera distribuzione delle perdite oltre il valore soglia.

Come evidenziato poco sopra, queste due nuove misure di rischio non sono esenti da difetti: vi è dunque una sorta di trade-off tra coerenza e semplicità di utilizzo. Una certa maneggevolezza in ambito operativo può ad esempio portare ad avere una misura di rischio che, in determinate circostanze, non soddisfi l'assioma di subadditività e generi perciò misurazioni paradossali.

2.3.3 *Expected Shortfall*

In letteratura si è cercato di costruire una misura di rischio che coniugasse contemporaneamente i pregi della TCE e della WCE e si è giunti ad una soluzione alternativa e più adeguata rispetto alle precedenti, rappresentata dalla misura nota come *Expected Shortfall (ES)*. Essa è interpretabile finanziariamente come la perdita media considerando tutte le perdite oltre un certo valore soglia rappresentato dal VaR.

L'*Expected Shortfall* viene chiamato anche *Conditional Value-at-Risk (CVaR)*, in quanto fornisce una stima della perdita attesa, tenendo in considerazione tutte le possibili perdite che superano la soglia del VaR²³. Esso è una misura di rischio più conservativa rispetto al VaR, dato che tiene in considerazione la forma della coda sinistra della distribuzione e pone quindi maggior attenzione alle perdite. Per questi motivi con Basilea III²⁴ l'*Expected Shortfall* ha sostituito

²³ È quindi una perdita attesa (valore atteso delle perdite) condizionata al superamento del VaR.

²⁴ Accordo stipulato nel 2010, come conseguenza della crisi finanziaria del 2007-2009 che ha rivelato una profonda inadeguatezza della regolamentazione finanziaria allora vigente. Con questo accordo si sono ad esempio rafforzati i requisiti patrimoniali delle

il precedente Value-at-Risk per quanto riguarda la disciplina sui rischi di mercato.

Seguendo la formulazione originale, ripresa da Acerbi e Tasche (2002), è ora possibile definire formalmente questa misura di rischio, al fine di comprenderne appieno la sua enorme funzionalità.

Data una distribuzione dei profitti e delle perdite X e specificato un determinato holding period e un livello di significatività $\alpha \in [0; 1]$, l'Expected Shortfall è definito come segue²⁵:

$$ES_{\alpha}(X) \stackrel{\text{def}}{=} -\frac{1}{\alpha} \left(\mathbb{E} \left[X I_{\{X \leq x^{(\alpha)}\}} \right] - x^{(\alpha)} (\Pr[X \leq x^{(\alpha)}] - \alpha) \right),$$

dove $x^{(\alpha)}$ equivale al $VarR_{\alpha}$.

Tale formula esprime un concetto semplice di cui si è già parlato precedentemente. In particolare il secondo addendo della somma algebrica all'interno della parentesi è interpretabile come la quantità da sottrarre al valor medio quando $X \leq x^{(\alpha)}$ ha probabilità superiore ad α . Quando invece $\Pr[X \leq x^{(\alpha)}] = \alpha$, come avviene normalmente nel caso di distribuzioni di probabilità continue, il secondo termine della somma algebrica si annulla e la formula dell' ES_{α} diventa coincidente con quella relativa alla TCE_{α} .

Per comprendere meglio l'effettiva semplicità di tale misura di rischio è utile fornire una rappresentazione equivalente alla precedente ma che presenta il vantaggio di una maggiore interpretabilità, rinunciando innanzitutto alla

banche ed è stata migliorata la gestione dei rischi al fine di migliorare la stabilità del sistema bancario. Basilea III ha avuto piena attuazione il primo gennaio 2019.

²⁵ $I_{\{X \leq x^{(\alpha)}\}}$ è la funzione indicatrice degli eventi che è pari a 1 se la relazione tra parentesi è soddisfatta, altrimenti è pari a 0.

definizione in termini di valore atteso. Indichiamo la funzione di densità di probabilità con $F(x) = \Pr(X \leq x)$ e introduciamo la funzione inversa di $F(x)$, $F^{-1}(p) = \inf\{x | F(x) \geq p\}$. È dimostrabile (Acerbi e Tasche, 2002) che l'Expected Shortfall può essere espresso come il valore atteso negativo di $F^{-1}(p)$ con un intervallo di confidenza $p \in (0; \alpha]$:

$$ES_{\alpha}(X) = -\frac{1}{\alpha} \int_0^{\alpha} F^{-1}(p) dp.$$

Questa formulazione dell' ES_{α} è nota in particolare per la sua trattabilità matematica che la rende appropriata per lo studio delle proprietà analitiche di questa misura di rischio. L'integrale definito permette di calcolare il valore medio della funzione inversa compresa tra 0 e α , ossia il valore medio delle perdite superiori alla soglia data dal VaR. La formula per il calcolo dell'Expected Shortfall (CVaR), con un livello di confidenza $1 - \alpha$ (che corrisponde al livello di rischio α), può anche essere riscritta nei seguenti termini:

$$CVaR_{1-\alpha}(X) = -\frac{1}{\alpha} \int_0^{\alpha} VaR_{1-t}(X) dt.$$

Così si percepisce bene il legame che sussiste tra il Value-at-Risk e l'Expected Shortfall.

Un'altra formulazione alternativa ma equivalente alle precedenti per indicare l'Expected Shortfall è stata introdotta di recente da Rockafellar e Uryasev (2001) ed è la seguente:

$$ES_{\alpha} = TCE_{\alpha} + (\lambda - 1)(TCE_{\alpha} - VaR_{\alpha}),$$

dove $\lambda \equiv \frac{\Pr[X \leq x^{\alpha}]}{\alpha} \geq 1$. Questa relazione è estremamente rilevante in quanto permette di mettere in evidenza che $ES_{\alpha} \geq TCE_{\alpha}$.

Nella Figura 2.5 si vuole mettere in luce quanto detto finora, rappresentando graficamente sia il VaR che l'Expected Shortfall al fine di maggiore chiarezza espositiva.

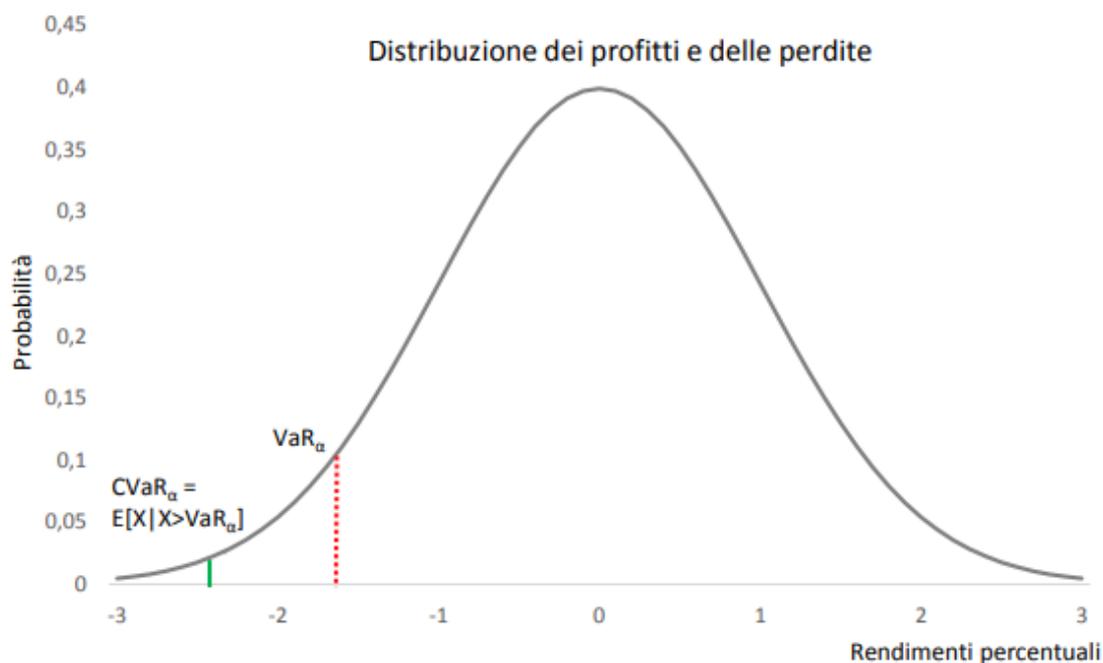


Figura 2.5 - L'Expected Shortfall come valore medio delle perdite che eccedono il VaR

L'Expected Shortfall risulta essere una misura di rischio applicabile ad ogni strumento finanziario e ad ogni fonte di rischio sottostante. Gode inoltre della proprietà di completezza, in quanto fornisce un numero riassuntivo della rischiosità globale del portafoglio esposto a differenti fonti di rischio. È inoltre una misura concettualmente semplice data la naturalezza e legittimità della domanda a cui l'ES fornisce una risposta. Infine, un'altra proprietà fondamentale che non si può assolutamente tralasciare è quella relativa alla

robustezza²⁶: l'Expected Shortfall, essendo una misura di rischio robusta, fornisce risultati sufficientemente precisi anche se si varia il livello di confidenza di qualche punto base. Quest'ultimo aspetto non è invece garantito nel caso in cui si applichino VaR, TCE o WCE.

Ora possiamo quindi ben comprendere i vantaggi dovuti dall'applicazione di tale strumento all'interno delle istituzioni finanziarie, vantaggi che rendono l'Expected Shortfall una misura di rischio sicuramente più appetibile rispetto alle precedenti. Per di più, come sostengono Acerbi e Tasche (2002), si ritiene che le banche possano facilmente passare da un sistema di risk management basato sul VaR ad uno basato sull'ES senza grandi sforzi computazionali; questo perché tutte le tecniche di calcolo del Value-at-Risk sviluppate negli anni recenti possono essere sviluppate senza eccessiva difficoltà anche al calcolo della misura di rischio coerente denominata Expected Shortfall.

2.4 Le misure di rischio entropiche

Sebbene le misure coerenti di cui si è parlato finora siano senza dubbio le più diffuse e conosciute, negli anni più recenti sono state introdotte molte altre misure di rischio, più o meno sofisticate e sempre coerenti, nel tentativo di fornire una sempre più elevata gamma di strumenti idonei alla misurazione del rischio nella moderna e sempre più complessa realtà finanziaria.

Ai fini del presente lavoro, verrà illustrata una particolare classe di misure, chiamate *misure di rischio entropiche*, ed in particolar modo verrà analizzato nel dettaglio l'*Entropic Value-at-Risk (EVAR)* la cui formulazione, come vedremo, rende evidente il fatto che esso è un metodo di calcolo innovativo ed originale

²⁶ In statistica si ha robustezza quando i risultati prodotti da un test statistico non variano in maniera considerevole se una o più ipotesi sulle quali è fondato il modello teorico non sono completamente rispettate.

per la misurazione del rischio di portafoglio. L'EVaR sarà dunque il nostro punto di partenza, o meglio dire la nostra funzione obiettivo, che dovrà essere minimizzata per risolvere il problema di ottimizzazione che verrà illustrato nei seguenti capitoli.

2.4.1 *La bioeconomia*

Per comprendere i concetti appena introdotti è prima di tutto necessario illustrare il contesto economico ma anche sociale e culturale in cui questa nuova classe di misure di rischio ha cominciato ad imporsi. Ezra Pound²⁷ sosteneva, infatti, che non si può comprendere un'epoca se non si approfondisce il sapere che più la caratterizza ovvero se non se ne intende appieno la specifica "cifra" culturale.

La storia contemporanea, in particolare nell'ultimo secolo, è stata caratterizzata da un grande impiego della *Tecnica*, intesa come il "saper fare" che mira al raggiungimento di un obiettivo o alla costruzione di un prodotto. Economia e tecnica economica e finanziaria sono diventati oramai degli strumenti fondamentali per l'umanità, ma allo stesso tempo essi si sono rivelati vincolanti in maniera alquanto rigorosa, poiché permeano ogni aspetto della realtà in cui viviamo. In un contesto sociale caratterizzato da discipline rigorose e da vincoli decisamente consolidati risulta utile riscoprire il pensiero di coloro che si sono collocati in un filone alternativo a quella che era la tecnica e la scienza dell'epoca.

²⁷ È stato un poeta, saggista e traduttore statunitense, che trascorse la maggior parte della sua vita in Italia. Fu uno dei protagonisti del modernismo e della poesia di inizio XX secolo.

Una delle più celebri menti che ha abbracciato in quegli anni idee economiche nettamente contrastanti con quelle classiche è Nicholas Georgescu-Roegen. Nato in Romania, egli emigrò in America quando la sua Patria divenne comunista e divenne celebre per avere iniziato ad approcciarsi all'economia in un modo decisamente nuovo ed originale. Georgescu fu innanzitutto economista, ma anche matematico, ecologo, scienziato e filosofo della scienza. Fu un grande innovatore e teorizzatore dell'incontro tra diverse discipline quali l'economia, la biologia e la fisica e soprattutto fu il fondatore di una nuova disciplina, la *bioeconomia*.

La bioeconomia è una teoria economica che è stata formulata negli anni '60 e che ha comportato un ripensamento dei fondamenti dell'economia neoclassica, basandosi sul linguaggio delle scienze fisiche e naturali. Il nucleo principale del pensiero di Georgescu è infatti la scoperta, in termini epistemologici e pratici, che nessuna scienza può evitare di tenere in considerazione l'ineluttabilità delle leggi della fisica, ed in particolare il secondo principio della termodinamica²⁸. I concetti che verranno descritti nelle seguenti pagine sono, dunque, strettamente legati al mondo della fisica e sono poi stati applicati a differenti ambiti, come quello economico. Come verrà illustrato di seguito, l'unione tra due discipline così apparentemente differenti tra loro, come la fisica e l'economia, ha portato alla costruzione di misure di rischio coerenti ed innovative che coniugano i pregi di misure quali il Value-at-Risk con i vantaggi derivanti dall'applicazione delle leggi naturali all'ottimizzazione di portafoglio.

Esistono varie formulazioni equivalenti del secondo principio della termodinamica, tutte volte ad enunciare l'irreversibilità degli eventi termodinamici. Se il primo principio della termodinamica si può formulare in

²⁸ La termodinamica è il settore della fisica che studia le trasformazioni dell'energia termica in energia meccanica e viceversa.

maniera generale dicendo che l'energia totale dell'Universo è costante, per il secondo principio questa energia non può trasformarsi liberamente da una forma ad un'altra, ma esistono delle limitazioni. Ad esempio l'energia termica non può passare da un corpo più freddo ad uno più caldo, o più precisamente la probabilità che questo accada è tendente allo zero. Inoltre la produzione di lavoro è sempre accompagnata da un riscaldamento e questo significa che nell'Universo vi è una tendenza spontanea al passaggio verso una forma di energia cosiddetta termica - dunque, il calore - che non è mai completamente ritrasformabile in un'altra forma di energia. In altre parole vi è la tendenza verso una forma "degradata" di energia, non più utilizzabile e così si intuisce il significato dell'irreversibilità degli eventi termodinamici di cui si è parlato poco sopra. In accordo con il secondo principio, quindi, si può affermare che nei sistemi fisici esiste una tendenza alla degradazione dell'energia e alla sua dispersione nell'ambiente. Questo concetto è strettamente collegato al fatto che ogni trasformazione fisica spontanea avviene soltanto in un verso determinato.

Tornando al pensiero economico di Georgescu, bisogna sottolineare che, mentre l'impostazione neoclassica si occupava dello sfruttamento delle risorse ambientali in un'ottica funzionale al benessere dei consumatori, la bioeconomia invece affronta la questione ecologica partendo dalla termodinamica.

Il secondo principio, brevemente descritto poco sopra, viene così esteso in ambito economico e si giunge inevitabilmente alla regola in base a cui in ogni processo produttivo si determina una diminuzione del potenziale d'uso dell'energia. Si arriva anche a definire inevitabile la dissipazione della materia nelle trasformazioni industriali e manifatturiere, data l'impossibilità del riciclo totale degli scarti e dei rifiuti.

La bioeconomia, partendo da queste basi, fornisce un utile apparato teorico per consentire una profonda comprensione della rilevante problematica che vede in opposizione il concetto di crescita economica con quello di sviluppo sostenibile.

Georgescu intuì i limiti delle teorie economiche precedenti, limiti che consistevano nella scarsità delle risorse e nell'impossibilità di eludere le leggi della fisica. Egli sostenne che "la scienza economica ha eliminato la dimensione ecologica dal suo orizzonte" e che questo l'ha ridotta ad un sapere astratto, virtuale, disgiunto dalla realtà della biosfera: era perciò necessario recuperare la dimensione "biologica" dell'economia.

2.4.2 L'entropia

Tornando al secondo principio della termodinamica, si noti che un sistema fisico tende spontaneamente a passare da uno stato iniziale più ordinato ad uno finale più disordinato. Per fare un esempio supponiamo di avere un recipiente contenente un gas libero di espandersi in un altro contenitore vuoto, collegato al primo per mezzo di una valvola: inizialmente tutte le molecole del gas si trovano nel primo contenitore, ma se la valvola viene aperta il gas tende spontaneamente a passare nel secondo contenitore e le sue molecole così si distribuiscono uniformemente all'interno dell'intero volume disponibile, rappresentato dai due contenitori. Una volta che tale processo è avvenuto è altamente improbabile che il sistema torni spontaneamente allo stato iniziale e in questo sta appunto l'irreversibilità del processo. Il processo, nel passare dallo stato iniziale a quello finale, ha comportato un aumento del grado di disordine complessivo del sistema, dato dalla disposizione delle molecole che, una volta aperta la valvola, tendono a distribuirsi uniformemente occupando tutto lo spazio disponibile. Generalizzando, si può dire che un sistema fisico isolato soggetto a trasformazioni spontanee tende verso il suo stato di massimo disordine.

Per chiarire quest'ultima affermazione bisogna fare riferimento ad una grandezza che è stata introdotta in fisica da Rudolph Clausius²⁹ nel 1864 e va sotto il nome di *entropia*. Essa esprime il grado di disordine di un sistema fisico ed in questi termini il secondo principio della termodinamica viene formulato nella maniera seguente: ogni trasformazione spontanea di un sistema fisico isolato è irreversibile e porta ad un aumento dell'entropia.

L'entropia è una funzione di stato e in quanto tale dipende solo dallo stato iniziale e da quello finale del sistema mentre è indipendente dal tipo di trasformazioni che il sistema subisce nel passare da uno stato all'altro.

In un sistema termodinamico non isolato che subisce una trasformazione, la variazione di entropia di un processo reversibile si definisce come il rapporto tra la quantità di calore scambiato dal sistema e la temperatura assoluta a cui avviene lo scambio, ovvero $\Delta S = \frac{Q}{T}$. L'unità di misura dell'entropia nel Sistema Internazionale è Joule/Kelvin. In un processo irreversibile invece, considerando sempre un sistema non isolato, la variazione di entropia è maggiore rispetto al rapporto tra quantità di calore scambiato e temperatura. Considerando poi un sistema isolato la situazione cambia in quanto non vi sono scambi di calore con l'ambiente esterno: in questo caso la variazione di entropia è nulla se il processo è reversibile ed è maggiore di zero se il processo è irreversibile.

Se si considera l'Universo come un sistema isolato termicamente, nel quale tutti gli scambi di calore con un eventuale ambiente esterno sono nulli e nel quale le trasformazioni spontanee sono irreversibili, allora il secondo principio della termodinamica stabilisce che l'entropia dell'Universo è in continuo aumento. Questo implica che l'energia totale dell'Universo, pur mantenendosi costante in accordo con il primo principio, va sempre più verso un grado di massima degradazione: in altre parole l'energia utile dell'Universo, ossia quella che può essere spontaneamente trasformata in lavoro, è in continua diminuzione mentre

²⁹ È stato un fisico e matematico tedesco e uno dei fondatori della termodinamica.

a causa degli attriti cresce l'ammontare complessivo di energia termica, altresì detta calore.

Georgescu nei suoi lavori ebbe l'intuizione di considerare come sistema fisico l'intero pianeta e giunse alla conclusione che un'economia volta alla sovrapproduzione ed allo sfruttamento intensivo delle risorse è termodinamicamente controproducente a causa dell'accumulo di entropia, ovvero di disordine che porta alla fine all'esaurimento delle risorse.

Il concetto di entropia, seppur è stato concepito in ambito fisico, è stato poi esteso ad altri ambiti non strettamente fisici tra cui la teoria dell'informazione³⁰, le scienze sociali ed in particolar modo l'economia.

Finora il concetto di entropia è stato strettamente collegato a quello di disordine ma ai fini del presente lavoro è possibile associarlo ad un concetto apparentemente diverso ma di fatto molto simile, ossia quello relativo alla mancanza di informazione associata all'incertezza. Questi due concetti infatti sono molto più simili di quanto apparentemente possa sembrare: essi sono due modi diversi di descrivere la stessa identica cosa. Si pensi ad esempio al modo in cui si reagisce quando si entra in una stanza molto ordinata oppure, invece, in una dove vi è un elevato grado di disordine. Nel primo caso non vi è difficoltà ad orientarsi e si prova una sensazione rassicurante; nel secondo caso risulta difficile orientarsi e si avverte invece una sensazione di inquietudine. Le misure di queste due grandezze hanno anche la stessa struttura matematica. Si può poi affermare che un sistema disordinato richiede molta più informazione per essere conosciuto rispetto ad un sistema ordinato. Allo stesso modo, se si sente la necessità di avere molta informazione su una determinata "cosa",

³⁰ La teoria dell'informazione è una disciplina dell'informatica e delle telecomunicazioni il cui oggetto è l'analisi e l'elaborazione su base matematica dei fenomeni relativi alla misurazione e alla trasmissione di informazioni su un canale fisico di comunicazione.

questo significa che questa “cosa” è caratterizzata da un elevato numero di configurazioni assunte ovvero da un elevato grado di disordine.

Il concetto di entropia è stato esteso all’ambito economico partendo da due idee essenziali: il *buon senso* con cui si prende una decisione e la *probabilità* che un certo evento incerto possa verificarsi. L’applicazione congiunta di una decisione di buon senso presa in concomitanza del manifestarsi di un evento incerto porta alla definizione dell’*entropia economica*.

Per comprendere le ragioni dell’inserimento di tali concetti in ambito economico si pensi che l’economia è assimilabile ad un gioco in cui i partecipanti scommettono qualcosa per ottenere delle vincite e alla base di queste scommesse vi è l’incertezza degli eventi, per i quali l’ammontare della vincita è inversamente proporzionale alla loro probabilità di accadimento. Risulta quindi necessario comprendere l’evoluzione dell’entropia economica per effettuare una scelta dettata dal buon senso e per evitare che “il gioco dell’economia cessi per eccessive perdite”.

Cerchiamo ora di spiegare cosa si intende per buon senso, concetto definito accuratamente nel XVIII secolo dal matematico e fisico svizzero Daniel Bernoulli. Secondo Bernoulli, in un gioco caratterizzato da eventi con probabilità molto basse e in cui la vincita è inversamente proporzionale alla probabilità, il premio da pagare per partecipare al gioco deve essere proporzionato non al valore assoluto della vincita ma al suo valore relativo, che è tanto maggiore quanto, a parità di vincita, risulta minore la ricchezza posseduta dal giocatore all’inizio del gioco. Il buon senso consiste dunque nel considerare la condizione soggettiva iniziale del giocatore, che condiziona la sua percezione della vincita e quindi la somma di denaro che è disposto a pagare per partecipare al gioco. In questo modo il gioco risulta essere equilibrato e non rischia di cessare per elevate perdite. Bernoulli intuì quindi

l'utilità marginale decrescente³¹ del denaro affermando come segue: "La determinazione del valore di un oggetto deve essere basata non sul suo prezzo, ma piuttosto sulla utilità che può procurare; non c'è dubbio che un guadagno di mille ducati ha più valore per un povero che per un ricco, nonostante entrambi guadagnino la stessa quantità".

Egli propose così un'opportuna una funzione di utilità logaritmica³² che, essendo concava, è coerente con l'idea dell'utilità marginale decrescente di cui si è parlato poco sopra. Per calcolare correttamente il premio da pagare per partecipare al gioco è necessario calcolare il valore atteso di questa funzione, tenendo conto delle probabilità di realizzazione delle singole vincite. Questo porta a stabilire un legame logaritmico tra l'ammontare dei premi da pagare e le vincite che possono riscuotersi, inversamente proporzionali alla probabilità degli eventi incerti. Una scelta di buon senso, in condizioni di incertezza, è perciò una misura fatta in scala logaritmica sul reciproco di una probabilità, ossia sul numero di tutte le possibili configurazioni³³ assunte dal sistema economico sul quale si sta scommettendo³⁴. In questi termini viene ricalcata

³¹ Si intende una situazione in cui la derivata prima della funzione di utilità diminuisce all'aumentare dell'ascissa nel piano cartesiano e perciò la derivata seconda è negativa. La funzione di utilità in questione è quindi concava. Dal punto di vista pratico ciò mette in luce come un incremento unitario di ricchezza apporta un beneficio maggiore a un individuo con una ricchezza iniziale inferiore.

³² La funzione di utilità proposta da Bernoulli è del tipo $u(x) = \ln x$.

³³ Per evidenziare il parallelismo con il mondo della fisica si noti che le possibili configurazioni assunte dal sistema economico sono i microstati possibili assunti da un sistema fisico.

³⁴ Supponiamo ad esempio che il valore atteso della funzione di utilità logaritmica sia il seguente: $\mathbb{E}[\ln x] = \sum_{n=1}^{\infty} \ln(2^n) \frac{1}{2^n}$. Questo gioco consiste nel lanciare ripetutamente una moneta finché non esce *Croce*, che dà luogo alla vincita. La vincita dipende dal numero di lanci che precedono l'uscita di *Croce*. Dunque si paga un certo premio e si vince 2^{k-1} , se la moneta è stata lanciata k volte quando compare *Croce* per la prima volta. Il termine $\frac{1}{2^n}$ è la probabilità della rispettiva vincita e l'argomento del logaritmo indica tutte le configurazioni possibili del sistema economico.

esattamente la definizione formale data da Ludwig Boltzmann³⁵ in merito al concetto di entropia termodinamica³⁶.

2.4.3 *Entropic Value-at-Risk*

Partendo da concetti strettamente legati alla fisica e alle scienze naturali si è così assistito negli ultimi anni ad un mutamento del panorama finanziario. Il concetto di entropia, dunque, seppur concepito nel mondo della fisica, è stato esteso alla teoria dell'informazione (Shannon, 1948) dove misura l'ammontare medio di informazione prodotta stocasticamente da una determinata fonte.

A tal proposito, sono state introdotte ulteriori estensioni e generalizzazioni di questa nozione di entropia e la più nota va sotto il nome di entropia di Rényi (Rényi, 1961). Essa quantifica la diversità, l'incertezza e l'imprevedibilità di un sistema. Rényi definì inoltre una gamma di misure che generalizzano la divergenza di Kullback-Leibler, chiamata anche entropia relativa, che misura la distanza che vi è tra due distribuzioni di probabilità. In altri termini descrive la perdita di informazione quando si considera erroneamente una distribuzione g mentre la distribuzione corretta da considerare è la f .

³⁵ È stato un fisico, matematico e filosofo austriaco del XIX secolo. È stato uno dei più grandi fisici teorici di tutti i tempi. La sua fama è dovuta alle ricerche in termodinamica e meccanica statistica (l'equazione fondamentale della teoria cinetica dei gas e il secondo principio della termodinamica).

³⁶ In meccanica statistica l'entropia serve ad ottenere informazioni macroscopiche su determinate grandezze, come pressione e temperatura, a partire dalle configurazioni microscopiche. Si immagina che a una certa condizione macroscopica di equilibrio del sistema (macrostatato o stato termodinamico del sistema) corrispondano diverse configurazioni microscopiche (stati dinamici o microstati, definiti solo se si conoscono posizione e velocità di tutte le molecole del sistema). Tali configurazioni microscopiche occupano un volume nello spazio delle fasi indicato con la lettera W . Allora possiamo definire l'entropia (S) secondo il principio di Boltzmann come: $S = k \ln W$, dove k è la costante di Boltzmann e W può essere meglio definita come la misura dell'insieme di tutte le possibili disposizioni (o probabilità a livello macroscopico) dei livelli molecolari e quindi il numero di stati totali accessibili al sistema ad una certa temperatura.

L'entropia relativa risulta quindi essere una nozione fondamentale ai fini del presente lavoro, in quanto la misura di rischio che si illustrerà di seguito è strettamente legata ai concetti che si sono descritti in questo paragrafo.

Si è già visto come l'Expected Shortfall, chiamato anche CVaR, sia una misura di rischio coerente che ha guadagnato negli ultimi anni molta popolarità, superando i più significativi difetti del Value-at-Risk. Un grande limite del CVaR, però, è la sua mancata efficienza computazionale: la grandezza e la complessità del problema di ottimizzazione è strettamente dipendente dalla grandezza del campione di dati a disposizione, in quanto un grande campione comporta un gran dispendio di tempo computazionale. Il bisogno di una misura di rischio coerente ed efficiente dal punto di vista computazionale è ancora più marcato quando tale misura è incorporata in un problema di programmazione stocastica che necessita di essere risolto frequentemente.

Una misura di rischio innovativa, entropica e coerente che presenta notevoli vantaggi computazionali è stata introdotta da Amir Ahmadi-Javid (2012) e va sotto il nome di *Entropic Value-at-Risk (EVaR)*. Dal nome si può già intuire il legame che questa misura ha con il VaR e con l'entropia relativa.

Prima di illustrare la relativa formalizzazione, è opportuno soffermarci sulla notazione che verrà usata di seguito. Innanzitutto va specificato che generalmente le misure entropiche dipendono dall'avversione al rischio dello specifico individuo e noi indicheremo il grado di avversione con il parametro z che per definizione è una quantità sempre positiva. Si può così ben comprendere l'importanza di tale parametro in quanto queste misure, *ceteris paribus*, forniscono risultati diversi a seconda dell'attitudine al rischio del singolo individuo.

Si indicherà poi con \mathbf{L}_{M^+} l'insieme di tutte le funzioni misurabili boreliane³⁷ $X: \Omega^{38} \rightarrow \mathbb{R}$, la cui funzione generatrice dei momenti³⁹ $M_X(z) = \mathbb{E}(e^{zX})$ esiste per ogni $z \geq 0$.

L'EVaR risulta essere il minor limite superiore ottenuto dalla disuguaglianza di Chernoff per il VaR e per il CVaR agli stessi livelli di confidenza, con $X \in \mathbf{L}_{M^+}$ e per ogni $\alpha \in (0; 1]$. Tale disuguaglianza è cruciale ai fini della costruzione di questa misura di rischio e verrà illustrata successivamente. Vale perciò la seguente relazione:

$$\text{VaR}_{1-\alpha}(X) \leq \text{CVaR}_{1-\alpha}(X) \leq \text{EVaR}_{1-\alpha}(X).$$

Si può così intuire come l'EVaR sia una misura più cauta rispetto al CVaR allo stesso livello di confidenza, poiché fornisce dei valori più elevati e quindi più prudenti. Ciò fa ben comprendere come un'istituzione finanziaria che adotti l'EVaR all'interno del sistema di risk management debba allocare più risorse e quindi più capitale per proteggersi del rischio. Questo è per di più un aspetto che deve essere tenuto in considerazione dalla singola impresa: infatti l'EVaR può essere giudicato meno attraente da imprese che vogliono allocare un ammontare di risorse quanto più contenuto possibile.

³⁷ La σ -algebra di Borel è la più piccola σ -algebra su un insieme dotato di struttura topologica che sia compatibile con la topologia stessa, ovvero che contenga tutti gli aperti della topologia. Lo spazio misurabile così definito prende il nome di spazio boreliano, gli insiemi contenuti nella σ -algebra di Borel sono detti insiemi boreliani e una misura definita su una σ -algebra di Borel è detta misura di Borel. La σ -algebra su un certo insieme invece è una famiglia di sottoinsiemi con proprietà di chiusura rispetto ad alcune operazioni insiemistiche.

³⁸ Ω è l'insieme di tutti gli eventi elementari.

³⁹ La funzione generatrice dei momenti viene usata in statistica per caratterizzare in modo astratto le variabili casuali permettendo da un lato di estrarne agevolmente alcuni parametri (come il valore atteso e la varianza) e dall'altro di confrontare due diverse variabili casuali e vedere il loro comportamento in condizioni limite.

Per ciascuna costante a e con $X \in \mathbf{L}_{M^+}$ la disuguaglianza di Chernoff è la seguente:

$$\Pr(X \geq a) = \Pr(e^{zX} \geq e^{za}) \leq e^{-za} M_X(z), \quad \forall z > 0.$$

Nella teoria della probabilità la disuguaglianza di Chernoff fornisce il limite superiore ed inferiore nella distribuzione di probabilità cumulata della somma di variabili casuali indipendenti. Questa disuguaglianza può assumere diverse formulazioni, a seconda delle ipotesi distributive poste alla base del modello e nel nostro caso si assume il caso più generale che prescinde dalla forma distributiva. La formulazione in questione viene ricavata analiticamente dalla disuguaglianza di Markov, che fornisce il limite superiore della probabilità che una variabile casuale sia maggiore o uguale ad una costante positiva, e in termini analitici è la seguente:

$$\Pr(X \geq a) \leq \frac{\mathbb{E}(X)}{a}, \quad \forall a > 0.$$

Markov riuscì così a legare il concetto di probabilità con quello di valore atteso e a fornire un limite superiore nella funzione di distribuzione di probabilità cumulata di una generica variabile casuale.

Si osservi ora che il termine $\Pr(X \geq a)$ può essere scritto in maniera equivalente come $\Pr(e^{zX} \geq e^{za})$ e applicando a quest'ultimo la relazione di Markov si ottiene la seguente disuguaglianza:

$$\Pr(X \geq a) = \Pr(e^{zX} \geq e^{za}) \leq \frac{\mathbb{E}(e^{zX})}{e^{za}}.$$

Si noti che il termine a destra della disuguaglianza coincide con $e^{-za} M_X(z)$, in quanto per definizione al numeratore compare la funzione generatrice dei momenti ed il termine al denominatore può essere portato al numeratore dopo

che si è invertito il segno dell'esponente. Con questi semplici passaggi matematici siamo così giunti alla formulazione della disuguaglianza di Chernoff partendo dalla disuguaglianza di Markov.

La relazione proposta da Chernoff è cruciale per la presenza al suo interno della funzione generatrice dei momenti di cui è parlato poco sopra. È proprio grazie a questa funzione che all'interno della formulazione vi sono racchiuse tutte le informazioni relative ai singoli momenti della distribuzione di probabilità. Questo risulta estremamente utile in quanto i momenti contengono importanti informazioni sulla distribuzione stessa.

Risolvendo l'equazione $e^{-za} M_X(z) = \alpha$ rispetto ad a con $\alpha \in (0, 1]$, si ottiene:

$$a_X(\alpha, z) = z^{-1} \ln\left(\frac{M_X(z)}{\alpha}\right).$$

Possiamo ora riscrivere la disuguaglianza di Chernoff come:

$$\Pr(X \geq a_X(\alpha, z)) \leq \alpha.$$

Si noti così che per ogni $z > 0$, $a_X(\alpha, z)$ è un limite superiore per $VaR_{1-\alpha}(X)$. Consideriamo ora il minor limite superiore di questo tipo come nuova misura di rischio che vincoli il $VaR_{1-\alpha}(X)$ servendosi della funzione generatrice dei momenti⁴⁰. A questo punto, effettuate queste premesse, possiamo definire formalmente l'Entropic Value-at-Risk.

⁴⁰ Risulta utile evidenziare la differenza tra la funzione generatrice dei momenti, di cui si è parlato precedentemente, e i momenti semplici, che possono essere derivati dalla funzione generatrice. Nel caso di distribuzioni discrete il momento semplice di origine m e ordine k di una variabile casuale è definito come: $\mu_{m,k} = \sum_{i=1}^n (x_i - m)^k p_i$.

L'EVaR di una variabile casuale $X \in \mathbf{L}_{M^+}$ con un livello di confidenza $(1 - \alpha)$ è definito come:

$$EVaR_{1-\alpha}(X) \stackrel{\text{def}}{=} \inf_{z>0} \{a_X(\alpha, z)\} = \inf_{z>0} \left\{ z^{-1} \ln \left(\frac{M_X(z)}{\alpha} \right) \right\} = \inf_{z>0} \left\{ z^{-1} \ln \left(\frac{\mathbb{E}(e^{zX})}{\alpha} \right) \right\}.$$

Il coefficiente di avversione al rischio z rende questa misura di rischio adattabile a tutte le tipologie di individui, a seconda della loro predisposizione al rischio. Si noti, ad esempio, che più basso è il valore di z più l'EVaR fornisce dei valori bassi, poiché nel caso in esame gli individui sono più disposti a rischiare e quindi percepiscono un rischio più contenuto. Il contrario avviene, invece, se viene posto un valore sufficientemente elevato di z , in quanto in questo caso la percezione soggettiva del rischio da parte degli individui è maggiore e quindi minore è la loro predisposizione a rischiare.

L'EVaR può essere anche descritto mediante la cosiddetta *rappresentazione duale*⁴¹ o robusta, che mette in luce la relazione tra questa misura e l'entropia relativa nella forma della divergenza di Kullback-Leibler. La rappresentazione duale dell'EVaR è la seguente:

$$EVaR_{1-\alpha}(X) = \sup_{Q \in \mathfrak{S}} E_Q(X),$$

dove $X \in \mathbf{L}_M$ ⁴² e \mathfrak{S} è l'insieme delle misure di probabilità e in termini formali è pari a:

$$\mathfrak{S} = \{Q \ll P: D_{KL}(Q||P) \leq -\ln \alpha\}.$$

$D_{KL}(Q||P)$ è l'entropia relativa di Q rispetto a P , ossia la divergenza di Kullback-Leibler. La rappresentazione duale dell'EVaR mette in luce, dunque, il legame

⁴¹ Per approfondimenti sulla rappresentazione duale si rimanda a quanto scritto in Ahmadi-Javid (2012).

⁴² Con \mathbf{L}_M si intende l'insieme di tutte le funzioni misurabili boreliane $X: \Omega \rightarrow \mathbb{R}$, la cui funzione generatrice dei momenti $M_X(z) = \mathbb{E}(e^{zX})$ esiste per ogni $z \in \mathbb{R}$.

di questa misura di rischio con la mancanza di informazione misurata dall'entropia relativa.

Va sottolineato poi che tale misura di rischio presenta notevoli vantaggi rispetto alle misure illustrate precedentemente. Innanzitutto, come è già stato accennato, un'ottimizzazione di portafoglio basata sull'approccio dell'EVaR presenta maggiori vantaggi computazionali rispetto ad un approccio basato sul CVaR, in particolare se si aumenta la numerosità campionaria.

L'EVaR è inoltre una misura coerente per ogni $\alpha \in (0; 1]$ e, a differenza del VaR e del CVaR, è una funzione strettamente monotona in tutto il suo dominio, rispetto all'insieme delle variabili casuali continue che la costituiscono. La ragione per cui il CVaR non gode della proprietà di forte monotonicità è che tale misura di rischio considera solo le perdite superiori alla soglia data dal VaR e non tiene minimamente in considerazione quelle minori. L'EVaR invece tiene conto simultaneamente sia delle perdite inferiori che di quelle superiori, focalizzandosi sull'intera distribuzione di probabilità dei rendimenti. Questo è il motivo per cui l'EVaR è sempre maggiore del CVaR allo stesso livello di confidenza.

È stato dimostrato, inoltre, che se due distribuzioni di probabilità hanno lo stesso EVaR a tutti i livelli di confidenza, allora sono perfettamente identiche dato che hanno la stessa funzione generatrice dei momenti. Questo è un ulteriore vantaggio che rende l'EVaR una misura di rischio sicuramente più "completa" ed affidabile rispetto ad altre misure come ad esempio il Value-at-Risk⁴³.

L'EVaR è infine cruciale ai fini della risoluzione di problemi di programmazione stocastica dove le variabili casuali seguono distribuzioni indipendenti ed arbitrarie oppure quando seguono un modello multifattoriale

⁴³ Uno dei problemi del Value-at-Risk è che si può avere lo stesso VaR a tutti i livelli di confidenza e nonostante ciò avere due diverse distribuzioni di probabilità.

lineare generalizzato⁴⁴: in questi casi infatti il problema risulta estremamente complesso e non può essere risolto efficacemente ed efficientemente con le altre misure di rischio viste in precedenza.

⁴⁴ I modelli lineari generalizzati (GLM) sono una generalizzazione del modello lineare nell'ambito della regressione lineare. Mentre nel modello lineare classico si ipotizza che la variabile dipendente (endogena) sia distribuita in modo normale, nell'ambito dei modelli lineari generalizzati la variabile endogena può essere distribuita come una qualsiasi variabile casuale della famiglia esponenziale come ad esempio la normale, la binomiale, la poissoniana, la gamma e la normale inversa. Se un modello presenta due o più variabili esplicative (esogene) allora prende il nome di modello multifattoriale.

CAPITOLO 3

UN MODELLO DI SELEZIONE DI PORTAFOGLIO

3.1 La funzione obiettivo

Nei primi due capitoli sono state date le basi per la specificazione di un modello di selezione di portafoglio. Nel corso di questo capitolo verrà concretamente costruito il modello di selezione oggetto del presente lavoro. Nel caso specifico si tratta di risolvere un problema di minimizzazione vincolata, in quanto si vuole trovare la combinazione di titoli in portafoglio che minimizzi la rischiosità complessiva del portafoglio stesso sotto determinati vincoli che illustreremo di seguito.

Le variabili decisionali sono le percentuali di capitale iniziale investite nei singoli titoli.

La funzione obiettivo è una relazione che nel caso in esame esprime la rischiosità complessiva del portafoglio e dev'essere quindi minimizzata; essa è funzione delle variabili decisionali. Nel precedente capitolo sono state illustrate le diverse misure di rischio ed i loro limiti e si è giunti alla conclusione che le misure di rischio coerenti soddisfano le proprietà ritenute desiderabili ai fini di una corretta ottimizzazione di portafoglio. Tra le varie misure coerenti è stata scelta come funzione obiettivo del problema di selezione l'Entropic Value-at-Risk, una nuova misura di rischio che presenta grandi vantaggi computazionali rispetto all'Expected Shortfall e che possiede un legame con il Value-at-Risk e con l'entropia relativa.

Si ricorda di seguito la formulazione dell'EVaR di una variabile casuale $X \in \mathbf{L}_{M^+}$ con un livello di confidenza $1 - \alpha$:

$$EVaR_{1-\alpha}(X) = \inf_{z>0} \left\{ z^{-1} \ln \left(\frac{\mathbb{E}(e^{zX})}{\alpha} \right) \right\}.$$

Il coefficiente di avversione al rischio z , per definizione sempre positivo, è cruciale in questa relazione in quanto, come sottolineato nel capitolo precedente, rende tale indice di rischiosità adattabile a tutte le tipologie di individui, a seconda della loro inclinazione al rischio.

L'EVaR, per ogni variabile casuale X che rappresenta il rendimento di un certo titolo o portafoglio, fornisce un numero esplicativo della rischiosità di quell'investimento. Il numero fornito in output dalla formula dipende anche da α , generalmente posto uguale a 1% o 5%; questa percentuale indica la probabilità di realizzazione di perdite più elevate rispetto a quelle calcolate tramite l'EVaR con un livello di confidenza $1 - \alpha$. In altre parole un livello di confidenza del 95%, e quindi un α pari a 5%, presuppone che l'EVaR misuri la rischiosità tenendo in considerazione tutte le potenziali perdite del portafoglio inferiori a quelle rappresentate dalla soglia α ; è perciò escluso il 5% di perdite situate nella coda sinistra della distribuzione di probabilità, che sono le perdite peggiori e quindi più elevate. Se invece α fosse una percentuale minore, ad esempio 1%, l'EVaR considererebbe il 99% di tutte le perdite realizzabili, mentre non verrebbero misurate le perdite più elevate conseguibili nell'1% dei casi; in questo caso la misura di rischio indicherebbe una rischiosità più elevata, in quanto si sta adottando un atteggiamento più prudente secondo il quale vengono considerati anche gli scenari più critici e meno probabili. Si può così comprendere l'importanza, ai fini di una buona specificazione del modello, dell'individuazione di un adeguato livello di confidenza; a seconda che si voglia adottare un atteggiamento più prudente oppure più audace si ritiene

quindi più consono attribuire ad α valori inferiori nel primo caso e invece più elevati nel secondo.

3.2 Applicazione della misura di rischio

Nel seguente paragrafo sarà presentata l'applicazione della misura di rischio appena descritta. È però necessario specificare prima la notazione che verrà utilizzata:

- X è la variabile casuale continua che modella il rendimento del portafoglio e X_i rappresenta il rendimento del portafoglio i -esimo;
- r_i è il tasso di rendimento del titolo i -esimo;
- $\hat{r}_i = \mathbb{E}(r_i)$ è il rendimento atteso del titolo i -esimo.

Il rendimento complessivo del portafoglio, ricordando quanto detto nel primo capitolo, può essere scritto come:

$$R_p = \sum_{i=1}^N x_i r_i.$$

Il valore atteso del rendimento del portafoglio sarà invece pari a:

$$\hat{R}_p = \mathbb{E}(R_p) = \sum_{i=1}^N x_i \hat{r}_i,$$

in cui il rendimento atteso del titolo i -esimo viene stimato usando i dati storici, ossia:

$$\hat{r}_i = \frac{1}{T} \sum_{t=1}^T r_{it},$$

dove T è il numero di osservazioni storiche prese in considerazione per la stima del rendimento medio di un certo titolo e r_{it} è il rendimento del titolo i -esimo all'istante temporale t .

Poiché nella funzione obiettivo in questione assume un ruolo fondamentale la funzione generatrice dei momenti, possiamo implementare quest'ultima nella seguente maniera, per poi passare alla riformulazione dell'EVaR:

$$\mathbb{E}(e^{zX}) = \sum_{i=1}^N p_i e^{zX_i} = \frac{\sum_{i=1}^N e^{z(x_i r_{it})}}{T},$$

dove p_i rappresenta la probabilità di realizzazione di X_i . Si noti che il rendimento del portafoglio, espresso dalla variabile casuale X , viene scritto come la media ponderata dei rendimenti medi stimati dei singoli titoli contenuti al suo interno e le stime di questi rendimenti sono ottenute, per ciascun titolo, dalla media dei valori storici dei rendimenti di quel titolo.

L'espressione campionaria dell'equazione relativa all'Entropic Value-at-Risk può dunque essere riscritta come:

$$EVaR_{1-\alpha}(X) = \inf_{z>0} \left\{ z^{-1} \ln \left(\frac{\sum_{i=1}^N e^{z(x_i r_{it})}}{T\alpha} \right) \right\}.$$

3.3 Vincoli e complessità computazionale

Come è già stato evidenziato precedentemente, ai fini di un'adeguata specificazione del modello è necessario introdurre dei vincoli nel problema di selezione di portafoglio. I vincoli che considereremo all'interno del problema di ottimizzazione affrontato nel presente lavoro sono i seguenti:

- *vincolo di bilancio*: questo vincolo impone che l'intero capitale sia investito nel portafoglio ed è formulato come:

$$\sum_{i=1}^N x_i = 1;$$

- *vincolo sul rendimento medio minimo*: questo vincolo impone il tasso di rendimento atteso minimo che il portafoglio in questione deve soddisfare, ossia:

$$\sum_{i=1}^N x_i \hat{r}_i \geq \pi,$$

dove π è il *target return*, ossia il rendimento desiderato che si vuole ottenere. Ponendo $\hat{R}_p = \sum_{i=1}^N x_i \hat{r}_i$, come indicato nel paragrafo 3.2, il vincolo sul rendimento medio minimo può essere riscritto come:

$$\hat{R}_p \geq \pi.$$

Si noti che questi primi due vincoli compaiono anche nel problema di selezione proposto da Markowitz, affrontato nel capitolo 1, perché sono alla base di un qualunque problema di ottimizzazione di portafoglio.

I successivi vincoli che illustreremo, invece, sono necessari a rendere il problema più coerente con la realtà finanziaria:

- *vincoli di cardinalità*: si introduce questa classe di vincoli per contenere indirettamente i costi di transazione limitando il numero di titoli da selezionare. Quando infatti il numero di assets in portafoglio è elevato sono alti anche i costi di transazione da sostenere. Un numero contenuto

di titoli è invece più economicamente sostenibile. Il vincolo di cardinalità può essere formulato come:

$$K_l \leq \sum_{i=1}^N z_i \leq K_u, \text{ con } 1 \leq K_l \leq K_u \leq N,$$

dove:

- K_l è il numero minimo di titoli che il portafoglio può contenere;
- K_u è il numero massimo di titoli che il portafoglio può contenere;
- N è il numero totale di titoli presi in considerazione come alternative d'investimento ai fini della creazione del portafoglio desiderato;
- z_i è una variabile decisionale binaria che assume valore 1 se il titolo i -esimo è stato selezionato, altrimenti assume valore 0;

È necessario evidenziare che i vincoli di cardinalità contengono variabili discrete, come ad esempio il numero di titoli, a differenza dei primi due vincoli presentati che sono espressi da variabili continue. La compresenza nel sistema di vincoli di variabili sia discrete che continue rende il problema in questione un problema di *programmazione mista-intera*; come verrà illustrato di seguito, questo comporta notevoli difficoltà computazionali e ai fini della risoluzione del problema risulta necessario usare particolari tecniche algoritmiche.

- *vincoli di bound*: con questi vincoli si vuole inoltre limitare superiormente ed inferiormente le percentuali d'investimento in ciascun titolo (*upper* e *lower bound*). In questo modo si cerca di evitare l'investimento in titoli per un ammontare di capitale irrisorio o al contrario eccessivo, se

rapportato all'ammontare totale di capitale investito. Il vincolo può essere formalizzato nella maniera seguente:

$$z_i l_i \leq x_i \leq z_i u_i, \text{ con } 0 \leq l_i \leq u_i \leq 1,$$

dove:

- l_i è la minima frazione d'investimento possibile nel titolo i -esimo;
- u_i è la massima frazione d'investimento possibile nel titolo i -esimo.

Inoltre, per rendere compatibile il vincolo di bound con il vincolo di cardinalità devono essere soddisfatte le due seguenti disuguaglianze:

$$l_i \leq \frac{1}{K_l} \text{ e } u_i \geq \frac{1}{K_u}.$$

Il vincolo di lower bound presentato vieta anche le vendite allo scoperto. La vendita allo scoperto, o short selling, può essere espressamente vietata dalle autorità competenti e per questi motivi questo vincolo appare sensato e per nulla privo di fondamento empirico;

- *vincoli d'integralità*: gli ultimi vincoli necessari per risolvere il problema di selezione di portafoglio proposto sono così formulati:

$$z_i(z_i - 1) = 0, \text{ con } i = 1, \dots, N.$$

Tali vincoli indicano che la variabile decisionale z_i è binaria e può assumere solo i due valori 0 e 1. Esso sostituisce la formulazione $z_i \in \{0,1\}$. Se z_i assume il valore 0, allora il titolo i -esimo non è presente

nel portafoglio; al contrario, se z_i assume il valore 1, allora il titolo i -esimo appartiene al portafoglio e quindi contribuisce in termini di rischio e rendimento al portafoglio complessivo. Si comprende così il legame sussistente tra i vincoli d'integralità, di cardinalità e di bound.

Il problema di selezione di portafoglio può essere ora specificato nella maniera seguente:

$$\begin{aligned} & \min_{X,Z} EVaR_{1-\alpha}(R_p) \\ \text{s. t. } & \begin{cases} \sum_{i=1}^N x_i = 1 \\ \hat{R}_p \geq \pi \\ K_l \leq \sum_{i=1}^N z_i \leq K_u \\ z_i l_i \leq x_i \leq z_i u_i \quad \forall i = 1, \dots, N \\ z_i(z_i - 1) = 0 \quad \forall i = 1, \dots, N \end{cases} \end{aligned}$$

Osservando il sistema di vincoli si può comprendere la complessità del problema di ottimizzazione in questione. Si tratta di un problema non lineare¹ e non convesso², in cui alcuni vincoli sono a variabili miste-interi. Si noti, infatti, che i vincoli di cardinalità e di integralità sono formulati in termini di variabili intere, a differenza degli altri che contengono variabili continue.

I vincoli a variabili miste-interi danno ai problemi di programmazione matematica per la selezione statica di portafogli azionari una valenza operativa maggiore di quella posseduta dai problemi classici, poiché rendono tali

¹ Il vincolo di integralità può essere ricondotto ad un'equazione di secondo grado e quindi non è lineare.

² Un problema di programmazione convessa è un problema di minimo in cui la funzione obiettivo è convessa e l'insieme ammissibile è un insieme convesso (oppure un problema di massimo in cui la funzione obiettivo è concava e l'insieme ammissibile è convesso). Una proprietà dei problemi convessi è che, o non esistono soluzioni o vi è soluzione globale, ma non possono esistere soluzioni locali. Il problema in questione, non essendo convesso, può ammettere invece diverse soluzioni locali.

problemi più coerenti con la realtà finanziaria. Questa categoria di vincoli ha però almeno due tipi di implicazioni non banali:

- il primo è relativo al fatto che verificare l'ammissibilità del sistema dei vincoli di questi problemi di programmazione matematica è, in generale, un problema *NP-completo*³;
- il secondo è relativo al fatto che risolvere questi problemi è, in generale, un problema *NP-hard*⁴.

Il problema che si affronta nella presente tesi è, come sottolineato precedentemente, non lineare, non convesso e misto-intero. I problemi di programmazione concava costituiscono una classe molto ampia di problemi non convessi, che include numerose classi di problemi molto complessi e

³ Tale concetto è studiato nella teoria della complessità computazionale. In particolare per complessità temporale si intende il tempo che un algoritmo, ossia un procedimento sistematico di calcolo, impiega ad essere eseguito in funzione della lunghezza della sequenza dei caratteri che formano l'input. Per ordinare gli algoritmi in base alla loro complessità temporale si utilizza l'ordine di grandezza, indicato con la notazione $O(f(n))$, dove $f(n)$ è una funzione definita sugli interi non negativi n . Quindi, se considerato un numero iniziale di input pari a n , la complessità di un algoritmo è $O(f(n))$ se il numero di operazioni per eseguirlo è un multiplo di $f(n)$. Si definisce poi un algoritmo in tempo polinomiale un algoritmo che ha un tempo di esecuzione limitato superiormente da un polinomio nella dimensione dell'input per l'algoritmo, ossia $T(n) = O(n^k)$ con k costante. Il valore $T(n)$ è definito come il tempo massimo impiegato su qualunque input di dimensione n . I problemi per i quali esiste un algoritmo in tempo polinomiale fanno parte della classe di complessità P (Polynomial time). Invece, per i problemi della classe NP (Non-deterministic Polynomial time), non sono attualmente noti algoritmi di complessità polinomiale che riescano a risolverli. I problemi *NP-completi* possono infine essere descritti come quei problemi che sicuramente non appartenerebbero a P se P fosse diverso da NP ; al contrario, se si dimostra che anche un solo problema *NP-completo* appartiene a P , allora si avrebbe equivalenza tra P e NP . I problemi *NP-completi* possono così essere definiti come quelli che meno probabilmente appartengono anche a P e sono i più difficili delle classi P e NP .

⁴ I problemi *NP-hard* (o *NP-difficili*) possono essere informalmente definiti come la classe dei problemi almeno difficili come i più difficili problemi delle classi di complessità P e NP .

difficili da risolvere. La difficoltà principale che si manifesta nella soluzione di problemi di tipo concavo risiede nel fatto che possono essere presenti molti punti di ottimo locale, che non sono punti di ottimo globale. In questi casi si complica notevolmente il quadro risolutivo e la ricerca delle soluzioni globali può divenire un problema di natura combinatoria, che può richiedere tempi di calcolo esponenzialmente crescenti con le dimensioni del problema. È quindi necessario trovare il metodo risolutivo che più si adatti a tali circostanze e che riesca a calcolare il punto di ottimo globale in un tempo ragionevole. A tali fini bisogna generalmente rinunciare ad una tecnica di calcolo esatta e nel presente lavoro utilizzeremo perciò due algoritmi approssimativi, che forniscono una “buona” soluzione al problema in tempi computazionali ragionevoli. Nel successivo capitolo verrà illustrato il funzionamento dei due algoritmi che ci permetteranno di risolvere il problema di ottimizzazione di portafoglio oggetto di questa tesi.

CAPITOLO 4

PARTICLE SWARM OPTIMIZATION

E

FIREWORKS ALGORITHM

4.1 Dalle euristiche alle metaeuristiche

La Particle Swarm Optimization (PSO) e il Fireworks Algorithm (FWA) sono i due algoritmi che verranno implementati per risolvere il problema di selezione di portafoglio affrontato nel presente lavoro.

Un algoritmo è un procedimento rigoroso finalizzato a risolvere un determinato problema mediante un numero finito di passi elementari, chiari e non ambigui. Gli algoritmi possono essere di due diverse tipologie: gli *algoritmi esatti*, detti anche *completi*, e gli *algoritmi approssimativi*, detti anche *euristici*.

Con riferimento ai problemi di ottimizzazione combinatoria¹, gli algoritmi completi consentono di risolverli in modo esatto, cioè di trovare una soluzione ammissibile corrispondente all'ottimo della funzione obiettivo tra tutte le soluzioni ammissibili. L'applicazione di metodi esatti non è però sempre possibile, a causa della complessità computazionale di certe tipologie di problemi, come i problemi NP-hard, per i quali non esiste alcun algoritmo completo che riesca a trovare la soluzione ottima in tempo polinomiale. In altre parole, un algoritmo completo, supposto che riesca a trovare la soluzione ottima di un problema NP-hard, la trova in tempi non ragionevoli.

Gli algoritmi euristici, invece, a differenza degli algoritmi esatti, permettono di trovare una soluzione approssimata al problema, sacrificando la ricerca della

¹ L'ottimizzazione combinatoria studia i problemi di ottimizzazione in cui l'insieme ammissibile è discreto ed è quindi definito in termini di strutture combinatorie.

soluzione ottima a favore di tempi computazionali ragionevoli. Essi consentono così di risolvere in poco tempo, seppure in maniera approssimata, anche problemi NP-hard molto complessi, come quello che si è illustrato nel precedente capitolo. Queste procedure di calcolo risultano essere più appropriate ai nostri fini, in quanto nella realtà risulta più utile disporre di un buon criterio d'investimento, anche se non ottimo, calcolandolo però in un tempo relativamente breve.

L'uomo è biologicamente indotto a ragionare secondo un approccio euristico: questo avviene perché la nostra mente, quando si trova ad affrontare un problema complesso, lo semplifica ignorando parte delle informazioni, al fine di risparmiare tempo e sforzi. Questo avviene quotidianamente ogniqualvolta si vogliono assumere certe decisioni o compiere certe azioni. Per comprendere questo concetto consideriamo ad esempio il processo che avviene all'interno del nostro cervello quando si vuole afferrare una palla al volo. Se il nostro cervello dovesse eseguire preventivamente tutti i calcoli alla base del processo fisico, noi non faremmo in tempo a mettere in atto le azioni necessarie per afferrare la palla. Inoltre, potrebbero comparire dei disturbi improvvisi, come una folata di vento, che invaliderebbero il modello che sta alla base del processo decisionale razionale. Per questi motivi un approccio euristico basato su una sequenza di osservazioni ed aggiornamenti progressivi della propria posizione rispetto alla traiettoria della palla risulta essere un metodo più efficace rispetto all'approccio tradizionale. In questo modo riusciamo a trovarci nelle immediate vicinanze della palla, quando essa sta per cadere a terra, così da poterla afferrare al volo. Sicuramente esistono metodi migliori, più eleganti e spettacolari, per raggiungere tale obiettivo, ma se avessimo seguito un approccio razionale non avremmo avuto il tempo di metterli in pratica.

Una possibile classificazione degli algoritmi euristici è la seguente:

- *algoritmi costruttivi*: sono orientati alla costruzione di una soluzione ammissibile e prevedono tre passi fondamentali:
 - *inizializzazione*: viene scelto un input di partenza, ossia una soluzione iniziale;
 - *selezione di un nuovo input*: sulla base di un criterio di selezione individuato, il nuovo input dev'essere aggregato alla soluzione parziale trovata al passo precedente;
 - *criterio di arresto*: se la soluzione trovata non è ammissibile allora si torna al secondo passo, altrimenti si arresta l'algoritmo.

Le soluzioni individuate con tecniche costruttive dipendono strettamente dalla scelta dell'input iniziale e del criterio di selezione dei nuovi input da aggiungere alle soluzioni parziali. Questi algoritmi sono pertanto veloci, ma spesso la soluzione trovata è meno precisa rispetto ad altre tipologie di algoritmi;

- *algoritmi di ricerca locale*, anche detti *migliorativi*: sono orientati al miglioramento di una soluzione ammissibile e perciò risultano essere generalmente più precisi degli algoritmi costruttivi. Essi partono da una soluzione ammissibile e cercano iterativamente di migliorarla effettuando semplici modifiche della soluzione corrente. Un algoritmo di ricerca locale termina quando non esistono più modifiche del tipo prescelto in grado di migliorare la soluzione corrente. Il funzionamento di un algoritmo migliorativo può essere descritto dai seguenti quattro passi:
 - *inizializzazione*, dove viene scelta la soluzione iniziale per innescare il processo di ricerca;

- *individuazione di un intorno* della soluzione corrente, sulla base di un criterio prestabilito che consenta di esplorare il “vicinato” di tale soluzione;
- *scelta di una nuova soluzione*, corrispondente alla soluzione migliore individuata nell’intorno scelto;
- *criterio di arresto*, secondo cui l’algoritmo viene fermato se la nuova soluzione individuata non migliora.

Le soluzioni individuate dagli algoritmi di ricerca locale rappresentano ottimi locali e dipendono strettamente dalla scelta della soluzione di partenza e del criterio per l’individuazione dell’intorno.

- *metaeuristiche*: il termine meta-euristica è composto da due parole di origine greca: euristica deriva al verbo *heuriskein* che significa “cercare”, mentre il prefisso *meta* significa “oltre, ad un livello superiore”. Si può così intuire che le metaeuristiche sono delle tecniche volte a migliorare la qualità delle soluzioni individuabili con gli algoritmi costruttivi e di ricerca locale. Questi algoritmi cercano di trovare una soluzione esplorando lo spazio di ricerca in modo efficace ed efficiente, servendosi di particolari tecniche che permettono all’algoritmo stesso di non ricadere in un punto di ottimo locale. Fanno parte delle metaeuristiche sia la Particle Swarm Optimization che il Fireworks Algorithm.

Nel successivo paragrafo affronteremo in dettaglio quest’ultima categoria di algoritmi approssimativi, le metaeuristiche, che ci torneranno estremamente utili per risolvere il problema di ottimizzazione impostato nel capitolo precedente.

4.1.1 Le metaeuristiche

Il termine *metaeuristica* è stato coniato inizialmente dall'informatico americano Fred Glover² e, prima di diventare ampiamente accettato e adottato dalla collettività, si parlava genericamente di *euristica moderna*.

Negli ultimi 30 anni le metaeuristiche hanno subito profondi cambiamenti, tanto che la stessa definizione di metaeuristica è stata modificata numerose volte nel corso degli anni. Ad oggi non esiste ancora una definizione univocamente condivisa dagli studiosi e di seguito ne forniamo una che è stata recentemente proposta da Sörensen e Glover nel 2013:

“A metaheuristic is a high-level problem-independent algorithmic framework that provides a set of guidelines or strategies to develop heuristic optimization algorithms. The term is also used to refer to a problem-specific implementation of a heuristic optimization algorithm according to the guidelines expressed in such a framework.”

In altre parole, mentre l'euristica è un algoritmo che fornisce in tempi ragionevoli una soluzione che è una buona approssimazione della soluzione ottima, la metaeuristica approfondisce invece la ricerca della soluzione ottima nelle zone più promettenti dello spazio delle soluzioni. Mediante l'uso delle metaeuristiche è inoltre possibile risolvere qualsiasi problema di ottimizzazione e risultano così essere delle metodologie di calcolo alquanto generali.

Prima di affrontare i concetti cruciali relativi all'ambito delle metaeuristiche è necessario introdurre brevemente la nozione di *Swarm Intelligence*, che verrà approfondita nel successivo paragrafo. In questi termini si intende una forma di

² F. Glover (1986), *Future Paths for Integer Programming and Links to Artificial Intelligence*, Computers and Operations Research, Vol. 13, No. 5, pp. 533-549.

intelligenza complessa posseduta da certe tipologie di popolazioni di individui; questi, se considerati singolarmente, sono dotati di un'intelligenza estremamente elementare ma, all'interno del gruppo di appartenenza, sono in grado di creare una forma di intelligenza evoluta, mediante lo scambio di informazioni con gli individui dello stesso gruppo e quindi grazie ad una forma di cooperazione. Ogni individuo, quindi, agisce sia a livello individuale per soddisfare i propri bisogni che a livello collettivo per soddisfare i bisogni del gruppo e in ultimo anche di sé stesso.

Ora si possono illustrare i due concetti essenziali alla base delle metaeuristiche: la *diversificazione*, che si riferisce all'esplorazione dello spazio di ricerca da parte del singolo individuo (exploration) e l'*intensificazione*, che si riferisce invece allo sfruttamento dell'esperienza di ricerca che è stata fatta da tutti gli individui del gruppo fino a quel momento (exploitation). Quindi, mentre la diversificazione è legata ad un comportamento individualista, lo sfruttamento è legato ad un comportamento sociale, che si estrinseca nell'utilizzo dell'esperienza di ricerca fornita dagli altri individui del gruppo di appartenenza. Questi concetti sono estendibili all'intero sistema sociale, in quanto è proprio grazie all'individualismo e alla socialità che si ha la crescita e l'evoluzione delle società. È poi fondamentale trovare il giusto equilibrio tra la diversificazione e l'intensificazione, da una parte per identificare velocemente le regioni promettenti nella ricerca della soluzione ottima, e dall'altra per evitare di sprecare tempo in zone già esplorate o poco promettenti.

Di seguito forniamo una breve panoramica delle metaeuristiche, classificandole in base alle caratteristiche che le differenziano:

- *nature-inspired vs non-nature inspired*. Questa distinzione si basa sull'origine dell'algoritmo. Alcuni, infatti, traggono ispirazione dai comportamenti osservati in natura, come ad esempio gli Algoritmi

Genetici, che si ispirano alle teorie evuzionistiche di Charles Darwin, e l'Ant Colony Optimization, che trae spunto dal comportamento delle colonie di formiche. Vi sono poi altri algoritmi che non hanno nulla a che fare con la natura, come ad esempio il Fireworks Algorithm, che si ispira all'esplosione dei fuochi d'artificio. Questa classificazione, però, non è molto significativa nella pratica, perché recentemente sono stati sviluppati algoritmi ibridi che potrebbero appartenere ad entrambe le categorie, come a nessuna e, inoltre, anche per le metaeuristiche non ibride, non è sempre chiara e immediata la loro classificazione;

- *population-based vs single point search*. La differenza tra queste due categorie si basa sul numero di soluzioni calcolate contemporaneamente. Gli algoritmi che analizzano soluzioni singole sono detti *trajectory methods*, in quanto descrivono una traiettoria nello spazio di ricerca nel tentativo di individuare la soluzione ottima. Le metaeuristiche *population-based*, invece, nel processo di ricerca della soluzione ottima, generano e gestiscono nello stesso momento un insieme di soluzioni, chiamato popolazione;
- *funzione obiettivo statica vs funzione obiettivo dinamica*. Mentre in alcuni algoritmi la funzione obiettivo rimane la stessa, in altri viene modificata durante la ricerca con lo scopo di "scappare" da eventuali punti di minimo locale;
- *una struttura di vicinato³ vs molteplici strutture di vicinato*. Alcune metaeuristiche si servono di una sola struttura di vicinato, intendendo così che la topologia del vicinato non cambia durante il processo di ricerca dell'ottimo. Altre, invece, modificano dinamicamente durante il

³ Il *vicinato* di un individuo (o soluzione) x è un determinato intorno di x , che viene esplorato dall'algoritmo nella ricerca di soluzioni migliori rispetto a quella corrente. La struttura di vicinato è la struttura, intendendo così la forma e la grandezza, dell'intorno preso in considerazione.

processo tale struttura topologica, con l'obiettivo di diversificare il più possibile la ricerca;

- *uso della memoria vs non uso della memoria.* Gli algoritmi che non si servono della memoria passata agiscono come un processo di Markov⁴, dove l'informazione usata per determinare la successiva mossa dipende esclusivamente dall'informazione corrente. Tra gli algoritmi che usano la memoria, invece, si possono distinguere quelli con memoria a breve termine e quelli con memoria a lungo termine; mentre i primi si avvalgono solo di informazioni recenti, i secondi utilizzano dei parametri sintetici relativi alla ricerca passata per potersi muovere anche in base alla memoria a lungo termine. Al giorno d'oggi l'uso della memoria è uno degli elementi fondamentale per poter implementare metaeuristiche efficaci ed efficienti.

4.2 Swarm Intelligence

Il termine *Swarm Intelligence*, traducibile come *intelligenza di sciame* è stato coniato per la prima volta nel 1988 da Gerardo Beni, Susan Hackwood e Jing Wang. Si intende così il comportamento collettivo che un gruppo di individui mette in atto, interagendo tra di loro e seguendo determinate regole. Mentre il singolo agente può essere considerato non intelligente, poiché da solo dispone di capacità limitate, la collettività di agenti nel suo insieme può dare vita ad una forma di intelligenza superiore, collettiva e quindi definita *Swarm Intelligence*.

Un'intelligenza di questo tipo è presente in natura ed è stata ampiamente studiata dai ricercatori soprattutto con riferimento al comportamento sociale

⁴ Un processo stocastico markoviano è un processo aleatorio dove la probabilità che determina il passaggio da uno stato attuale a quello successivo è determinata solo dallo stato attuale e non dagli stati passati.

proprio di alcuni insetti, come le formiche, le termiti, le api e le vespe e anche di altri animali, come ad esempio gli uccelli e i pesci. Dall'analisi effettuata in merito al comportamento tenuto da certi insetti rispetto alla collettività è emerso che, questi animali, mentre singolarmente si limitano a compiere azioni elementari, a livello di gruppo riescono a costruire un'organizzazione collettiva molto complessa e di alto livello. La Swarm Intelligence è quindi un'intelligenza complessa posseduta da certe tipologie di popolazioni di individui e si basa su alcune caratteristiche essenziali: l'intelligenza di ciascun individuo della popolazione è limitata e vi è auto-organizzazione, il che significa che non vi è un coordinatore centrale che supervisiona l'intero processo, poiché ogni individuo della popolazione interagisce, direttamente o indirettamente, con gli individui dell'ambiente. È la comunicazione tra questi individui, e quindi la cooperazione tra di essi, che porta ad una forma di intelligenza più evoluta, ben oltre le capacità di ogni singolo soggetto.

A scopo esemplificativo si pensi al comportamento che le formiche tengono in natura con lo scopo di ricercare il cibo. Ogni formica, nel proprio cammino, deposita sul terreno una sostanza chimica, il feromone. La scelta del percorso da seguire alla ricerca del cibo è guidata dall'intensità del feromone; più esso è intenso, più è probabile che le altre formiche scelgano quella direzione. Il feromone però evapora nel tempo e quindi rimangono più "marcati" i percorsi scelti più di frequente. Alla fine del processo tutte le formiche convergeranno verso lo stesso percorso, che risulta essere il percorso ottimo, o una buona approssimazione di quest'ultimo e quindi più breve, che le conduce dalla fonte di cibo al formicaio. Si può così notare come le formiche, animali dotati di un'intelligenza estremamente elementare se presi singolarmente, riescano, mediante la collaborazione a livello di gruppo, a risolvere un problema di ottimizzazione complesso. Osservando quindi il comportamento in natura di certe tipologie di animali, sono state implementate una serie di metaeuristiche

basate sulla Swarm Intelligence, come ad esempio la Particle Swarm Optimization e il Fireworks Algorithm, che tratteremo di seguito.

Per concludere proponiamo un breve elenco esemplificativo e non esaustivo degli algoritmi basati sull'intelligenza di sciame, suddivisi in due classi a seconda che siano nature-inspired o non-nature inspired:

- nature-inspired: vi è ad esempio la Particle Swarm Optimization (PSO), l'Ant Colony Optimization (ACO), l'Artificial Bee Colony (ABC), il Bacterial Foraging Optimization (BFO), il Bat Algorithm (BA), il Firefly Algorithm (FA);
- non-nature inspired: ad esempio il Fireworks Algorithm (FWA), il Brain Storm Optimization (BSO), Magnetic Optimization Algorithm (MOA).

4.3 Particle Swarm Optimization

La Particle Swarm Optimization (PSO) è una metaeuristica di ispirazione biologica e population-based, introdotta nel 1995 da James Kennedy e Russel Eberhart⁵ e sviluppata ampiamente dagli stessi nel 2001⁶. La PSO appartiene alla classe degli algoritmi basati sulla Swarm Intelligence e si ispira al comportamento sociale degli stormi di uccelli e dei banchi di pesci. In natura vi è una forte cooperazione tra i volatili di uno stormo e tra i pesci di un banco. Questi animali riescono in gruppo a perseguire ottimamente degli obiettivi cruciali per la loro esistenza, come la ricerca del cibo e la fuga dai predatori. A

⁵ J. Kennedy, R. Eberhart, *Particle Swarm Optimization*, Proc. IEEE International Conference on Neural Networks (Perth, Australia), IEEE Service Centre, Vol. 4, pp. 1942-1948, Piscataway, NJ, 1995.

⁶ J. Kennedy, R. Eberhart, *Swarm Intelligence*, Morgan Kaufmann of Academic Press, New York, 2001.

tali fini, ogni individuo dello sciame⁷ si serve di informazioni relative sia alla propria velocità e posizione che alla velocità e alla direzione assunta dai membri appartenenti all'intera popolazione, riuscendo così a calibrare il proprio comportamento a vantaggio proprio e dell'intero sciame. Ciascun volatile all'interno dello stormo deve seguire determinate e rigorose regole finalizzate alla propria sopravvivenza e a quella dell'intero gruppo: deve seguire i propri vicini ed evitare di urtarli e inoltre deve, in generale, rimanere nello stormo. In ogni caso, un volatile che scorga una fonte di cibo può allontanarsi dal gruppo per raggiungerla, mettendo così in atto un comportamento individualista, oppure può rimanere all'interno del gruppo, dando luogo ad un comportamento sociale. Questi due aspetti contrapposti, se adeguatamente bilanciati, danno origine ad una efficace strategia di ricerca. L'*esplorazione*, legata all'individualità del singolo individuo che ricerca la soluzione ottima tra tutte quelle ammissibili, dev'essere bilanciata con lo *sfruttamento*, che è connesso alla socialità ovvero alla solidarietà tra gli individui del gruppo che comunicano tra loro le soluzioni identificate, al fine di trovare la soluzione ottima.

Nel caso di un problema di ottimizzazione, la popolazione di individui consiste in un determinato numero di particelle, che nel caso specifico identificano ognuna un portafoglio azionario. Ciascuna particella rappresenta una possibile candidata soluzione espressa in termini di vettore di variabili ed ha associato un valore che funge da indice di qualità di quella determinata soluzione; questo valore è misurato mediante un'opportuna funzione di fitness, ossia la funzione obiettivo che nel nostro caso è funzione dell'Entropic Value-at-Risk. Nell'applicazione di questo algoritmo, quindi, le particelle si spostano secondo opportuni criteri nello spazio di ricerca ammissibile al fine di trovare l'ottimo

⁷ Da qui in avanti useremo il termine *sciame* per indicare una collettività di individui omogenea e con una forte cooperazione interna, come avviene nel caso degli stormi di uccelli e dei banchi di pesci.

della funzione obiettivo. In questo caso si tratta di un problema di minimizzazione, ossia:

$$\min f(x), \text{ con } x \in \mathbb{R}^n,$$

dove $f: \mathbb{R}^n \rightarrow \mathbb{R}$ è la funzione obiettivo del problema di minimizzazione.

Ogni particella è caratterizzata da tre vettori, ciascuno di dimensione n , dove n sono le dimensioni dello spazio di ricerca. Se si considera una popolazione formata da M particelle in uno spazio multidimensionale, alla k -esima iterazione i tre vettori della particella i -esima, con $i \in \{1, \dots, M\}$, risultano essere i seguenti:

- $X_i^k \in \mathbb{R}^n$ è la posizione della particella i -esima alla k -esima iterazione, rappresentata da un vettore n -dimensionale:

$$X_i^k = (x_{i,1}^k, x_{i,2}^k, \dots, x_{i,n}^k);$$

- $V_i^k \in \mathbb{R}^n$ è la velocità della particella i -esima alla k -esima iterazione, rappresentata da un vettore n -dimensionale:

$$V_i^k = (v_{i,1}^k, v_{i,2}^k, \dots, v_{i,n}^k);$$

- $P_i \in \mathbb{R}^n$ è la migliore posizione in cui si è trovata la particella finora:

$$P_i = (p_{i,1}, p_{i,2}, \dots, p_{i,n});$$

- $G \in \mathbb{R}^n$ è la migliore posizione in cui tutte le particelle dello sciame si sono trovate finora:

$$G = (g_1, g_2, \dots, g_n).$$

Le M particelle sono inizialmente posizionate casualmente all'interno dello spazio delle soluzioni, in modo da assicurare una buona esplorazione dello spazio di ricerca. Ciascuna particella valuta ad ogni spostamento il proprio valore di ottimalità, determinato dalla funzione di fitness e corrispondente a $f(X_i^k)$. Nel caso in esame, poiché la funzione obiettivo misura il rischio del portafoglio⁸, la migliore posizione di ogni particella sarà quella associata al valore più basso della funzione di fitness. Ad ogni iterazione quindi la PSO ricalcola il valore di ottimalità per ciascuna particella e tiene in considerazione la migliore posizione raggiunta da ciascuna di esse, corrispondente al valore $f(P_i)$, chiamato *personal best*. Nella ricerca dell'ottimo l'algoritmo tiene in considerazione anche la migliore posizione che lo sciame di particelle ha raggiunto nel complesso durante il processo iterativo, identificata con la lettera G e corrispondente al valore di ottimalità $f(G)$, chiamato *global best*. Ciascuna particella dello sciame determina il proprio successivo movimento considerando la propria posizione corrente ma anche la propria migliore posizione passata e la migliore posizione passata dell'intero sciame. Questo permette all'algoritmo di esplorare lo spazio delle soluzioni ammissibili secondo dei criteri logici e ben definiti e di raggiungere in un ragionevole lasso di tempo una soluzione molto prossima rispetto al punto di ottimo del problema di ottimizzazione. Risulta evidente come questo sia un chiaro esempio di comportamento sociale, in quanto il comportamento individualista di ciascuna particella viene bilanciato dalla socialità, dettata da una forma di cooperazione a livello di sciame che porta all'identificazione della migliore soluzione possibile per i singoli individui e per il gruppo nel suo complesso.

Applicando la PSO al problema di selezione di portafoglio oggetto del presente lavoro, è necessaria la seguente specificazione. Si è detto precedentemente che

⁸ In realtà, come verrà meglio illustrato nel successivo capitolo, la funzione obiettivo è comprensiva sia della misura di rischio sia della funzione di penalità, che comprende le violazioni dei vincoli.

le singole particelle sono nel caso in esame portafogli azionari con una determinata composizione di titoli al loro interno. In questi termini il vettore X_i^k risulta essere il vettore composto dai pesi dei vari titoli azionari all'interno del portafoglio i -esimo e alla k -esima iterazione. Di conseguenza la migliore posizione passata di ciascuna particella, P_i , non è altro che il vettore dei pesi dei titoli associato al più basso valore della funzione di fitness, che nel caso specifico misura la rischiosità, calcolata per l' i -esimo portafoglio.

4.4 La struttura della PSO

La PSO ha subito negli anni vari adattamenti e progressi, in particolare per quanto riguarda l'equazione della velocità V_i^k e la struttura topologica del vicinato. Di seguito si propongono la versione originale di Kennedy e Eberhart del 1995 ed alcune successive modifiche.

4.4.1 L'algoritmo originale

L'algoritmo originale della PSO è strutturato nei seguenti passaggi:

1. Generazione iniziale della popolazione di particelle componenti lo sciame, attribuendo a ciascuna di esse posizione e velocità casuali.
2. Avvio ciclo:
 - a. Determinare il valore della funzione di fitness $f(X_i^k)$ relativa alla posizione corrente X_i^k di ciascuna particella;
 - b. Confrontare il valore della funzione di fitness di ogni particella con il proprio *personal best*. Se $f(X_i^k)$ è migliore di $f(P_i)$, allora si aggiorna il *personal best* della particella i -esima uguagliandolo alla

posizione nuova di tale particella. Nel caso in esame, poiché si tratta di un problema di minimizzazione, se $f(X_i^k) < f(P_i)$, allora bisogna impostare $P_i = X_i^k$;

- c. Identificare con la lettera G la particella che, tra quelle del vicinato⁹ di ogni particella, ha ottenuto il miglior valore della funzione di fitness, ossia il *global best*. Nel caso in questione si tratta della particella con il valore più basso della funzione obiettivo;
- d. Determinare la nuova posizione X_i^{k+1} e velocità V_i^{k+1} di ogni particella utilizzando le seguenti equazioni:

$$\begin{cases} V_i^{k+1} = V_i^k + U(0, \phi_1) \otimes (P_i - X_i^k) + U(0, \phi_2) \otimes (G - X_i^k) \\ X_i^{k+1} = X_i^k + V_i^{k+1} \end{cases},$$

dove:

- $U(0, \phi_i)$, con $i = \{1,2\}$, è un vettore di numeri casuali uniformemente distribuiti nell'intervallo $[0, \phi_i]$;
 - ϕ_1 e ϕ_2 sono i coefficienti di accelerazione, in particolare ϕ_1 esprime l'effetto che produce la migliore direzione individuale (coefficiente di accelerazione cognitivo), mentre ϕ_2 esprime l'effetto che produce la migliore direzione collettiva (coefficiente di accelerazione sociale);
 - \otimes è il prodotto tensoriale che di fatto è un moltiplicatore vettoriale;
- e. Se si soddisfa un predeterminato criterio di arresto, che può essere ad esempio un soddisfacente valore della funzione di fitness o un limite massimo di iterazioni, si passa al punto 3, altrimenti si ritorna all'inizio del punto 2 e, dopo aver aggiornato il numero

⁹ La struttura del vicinato può variare a seconda della topologia scelta. Si analizzerà questo aspetto nel sottoparagrafo 4.4.3.

delle iterazioni, si ripete il processo fino a quando non viene soddisfatto il criterio di arresto;

3. Fine del ciclo.

L'algorithmo appena descritto può essere rappresentato schematicamente dal seguente diagramma di flusso:

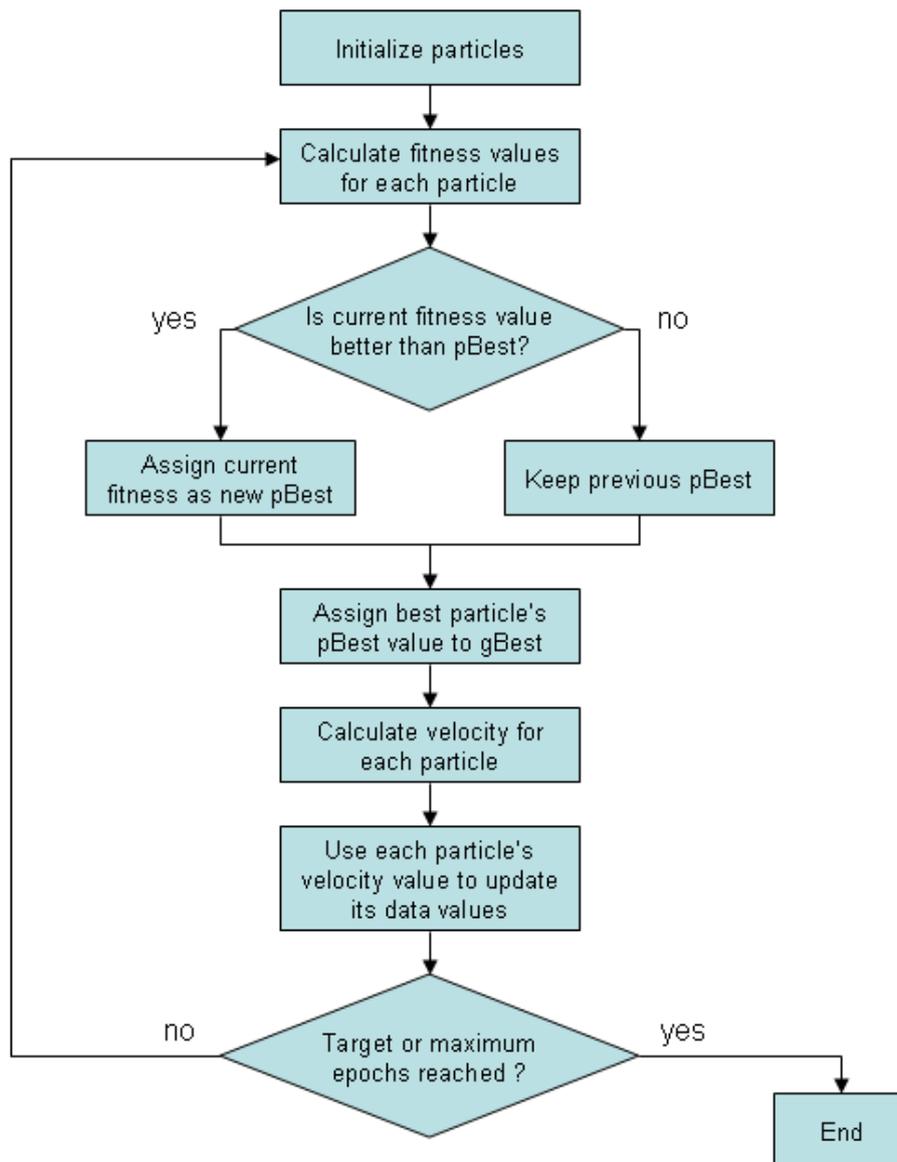


Figura 4.1 - Diagramma di flusso della PSO

Il punto più cruciale dell'intero processo è il calcolo della velocità V_i^{k+1} che serve a far spostare ogni particella nello spazio n -dimensionale dalla posizione X_i^k , relativa alla k -esima iterazione, alla posizione X_i^{k+1} , relativa all'iterazione $(k + 1)$.

L'equazione descritta poco sopra è costituita da tre addendi:

- V_i^k rappresenta la velocità corrente della particella i -esima;
- $U(0, \phi_1) \otimes (P_i - X_i^k)$ corrisponde al prodotto tra una componente aleatoria di cui si è parlato poco sopra e la differenza tra la migliore posizione passata dell' i -esima particella e quella attuale. Questo termine è la cosiddetta *componente cognitiva* e rappresenta la tendenza di ciascuna particella ad essere attratta verso la propria migliore posizione;
- $U(0, \phi_2) \otimes (G - X_i^k)$ rappresenta invece la parte dell'equazione legata alla cosiddetta *componente sociale*. $(G - X_i^k)$ è la differenza tra la migliore posizione passata globale di tutte le particelle appartenenti a un certo vicinato e la posizione corrente della particella i -esima. Questo termine rappresenta dunque la tendenza di ciascuna particella ad essere attratta verso la migliore posizione fra tutte quelle visitate da tutte le particelle del vicinato.

Le componenti cognitive e sociali sono di fatto delle correzioni del moto delle particelle, che spingono queste nella direzione dei punti di personal best e global best trovati fino a quel momento. All'interno del secondo e del terzo addendo compare una quantità aleatoria, corrispondente ad un vettore di numeri casuali; l'algoritmo cerca così di disturbare il moto della particella, in modo da spingerla ad approfondire la ricerca nelle zone limitrofe al personal best ed al global best.

I coefficienti di accelerazione ϕ_1 e ϕ_2 influenzano in maniera determinante l'intensità delle forze che attraggono ciascuna particella verso P_i e G . È quindi

molto importante stabilire dei corretti valori per questi due parametri, onde evitare un aumento incontrollato della velocità delle particelle e quindi l'instabilità dell'intero processo. Nei primi studi sulla Particle Swarm Optimization vennero utilizzati i valori di ϕ_1 e ϕ_2 pari a 2, che però non garantivano la stabilità del modello, in quanto si potevano verificare aumenti eccessivi della velocità. Si propose quindi un metodo di controllo della velocità, che limita questa all'interno di un determinato range di valori $[-V_{max}, +V_{max}]$. Sebbene si sia riusciti così ad evitare l'instabilità del processo, la scelta di V_{max} è alquanto complessa e non esistono regole definite per il suo settaggio. Per ovviare al problema sono state proposti altri metodi per controllare la velocità delle particelle, uno dei quali verrà illustrato nelle successive pagine.

Ai fini della corretta implementazione dell'algoritmo è poi necessario fissare un successivo parametro, relativo al numero di particelle componenti lo sciame. Un numero eccessivamente elevato di particelle comporterebbe notevoli sforzi computazionali, mentre un numero troppo basso ridurrebbe l'efficacia dell'algoritmo nella ricerca della soluzione ottima, in quanto sarebbe minore l'esplorazione dello spazio di ricerca; la prassi è perciò quella di generare uno sciame composto da un numero di componenti compreso tra 20 e 50 (Blackwell *et al.*, 2007).

4.4.2 Le modifiche alla PSO

4.4.2.1 Peso d'inerzia

Una prima modifica all'algoritmo generale della PSO fu proposta da Shi e Eberhart nel 1998 e va sotto il nome di *Inertia Weight*. L'equazione che definisce il valore della velocità futura di ciascuna particella diventa così la seguente:

$$\begin{cases} V_i^{k+1} = \omega V_i^k + U(0, \phi_1) \otimes (P_i - X_i^k) + U(0, \phi_2) \otimes (G - X_i^k) \\ X_i^{k+1} = X_i^k + V_i^{k+1} \end{cases},$$

dove ω è il peso d'inerzia.

Con questa modifica si è cercato di migliorare la potenzialità risolutiva dell'algoritmo; ponendo un freno inerziale sulla velocità iniziale di ogni singola particella, infatti, si dona maggiore stabilità al modello che altrimenti raggiungerebbe facilmente una condizione di seria instabilità.

Un'altra funzione molto importante del coefficiente inerziale è quella di bilanciare i due fondamentali comportamenti delle particelle, ossia lo sfruttamento e l'esplorazione. Lo sfruttamento è determinato dagli ultimi due addendi dell'equazione, mentre il comportamento esplorativo è specificato dalla componente ωV_i^k . Questa componente inerziale ha estrema rilevanza nel modello, in quanto, all'aumentare del valore attribuito a ω , aumenta l'influenza della velocità corrente su quella futura. Al contrario, un peso d'inerzia basso rende la velocità futura più dipendente dalla migliore posizione passata della singola particella e dell'intero vicinato.

Un criterio che permette di sfruttare il peso d'inerzia in modo razionale consiste nell'attribuirgli un valore decrescente all'aumentare del numero di iterazioni. Nella fase iniziale del processo di ricerca, infatti, risulta utile una maggiore esplorazione dello spazio n -dimensionale e quindi un valore elevato di ω . Successivamente, diminuendo il valore del peso d'inerzia, si predilige l'azione di sfruttamento consentendo all'algoritmo di approfondire la ricerca della soluzione ottima in una determinata regione dello spazio considerata "promettente". Si riporta di seguito una delle possibili formulazioni per determinare il valore del coefficiente inerziale all'aumentare delle iterazioni dell'algoritmo, utilizzato in Corazza *et al.* (2013):

$$\omega^k = \omega_{max} - \frac{\omega_{max} - \omega_{min}}{K} k,$$

dove:

- ω_{max} è un prefissato valore massimo per il peso d'inerzia;
- ω_{min} è un prefissato valore minimo per il peso d'inerzia;
- K è il numero massimo di iterazioni consentite;
- k rappresenta il numero dell'iterazione attuale.

La prassi è quella di attribuire a ω_{max} e ω_{min} rispettivamente i valori 0.9 e 0.4. Inoltre, per agevolare la convergenza verso l'ottimo del problema, deve essere rispettata la seguente relazione tra il peso d'inerzia e i due coefficienti di accelerazione ϕ_1 e ϕ_2 :

$$\frac{\phi_1 + \phi_2}{2} - 1 < \omega.$$

Il corretto settaggio del peso d'inerzia e dei coefficienti di accelerazione rende quindi la PSO più stabile rispetto al modello originale, tanto che la convergenza all'ottimo del problema è assicurata anche nel caso in cui venga assegnato un valore alto al parametro V_{max} oppure questo non venga neanche settato.

4.4.2.2 *Constriction coefficients*

Un'altra variante della PSO molto utilizzata sfrutta un metodo di limitazione dei coefficienti introdotto da Clerc e Kennedy nel 2002. In questo modo si riesce a prevenire l'instabilità del modello a causa di un'eccessiva velocità e viene garantita la convergenza delle particelle verso il punto di ottimo, risultando così superfluo il settaggio del parametro V_{max} . Il calcolo della velocità futura diventa ora il seguente:

$$\begin{cases} V_i^{k+1} = \chi \left(V_i^k + U(0, \phi_1) \otimes (P_i - X_i^k) + U(0, \phi_2) \otimes (G - X_i^k) \right) \\ X_i^{k+1} = X_i^k + V_i^{k+1} \end{cases},$$

dove:

- $\chi = \frac{2}{\phi - 2 + \sqrt{\phi^2 - 4\phi}}$;
- $\phi = \phi_1 + \phi_2$, con $\phi > 4$.

Clerc e Kennedy nella prassi hanno posto ϕ pari a 4.1, con $\phi_1 = \phi_2 = 2.05$; da ciò ne consegue un valore del moltiplicatore χ pari approssimativamente a 0.7298. A differenza di quanto si è visto in precedenza con riferimento al peso d'inerzia, il coefficiente di costrizione χ moltiplica tutte le componenti dell'equazione relativa alla velocità futura. Di conseguenza, mentre la velocità corrente è moltiplicata per 0.7298, gli ultimi due addendi, che esprimono lo sfruttamento dello spazio di ricerca, vengono moltiplicati per un valore casuale appartenente all'intervallo di valori equiprobabili $[0, 1.49618]$ ¹⁰. Si noti poi che le due varianti della PSO viste finora possano coincidere se si pone $\omega = \chi$ e $\phi_i = \chi\phi_i$; in questo modo nella PSO con il peso d'inerzia il valore del coefficiente inerziale diventa pari a 0.7298 e i due coefficienti di accelerazione diventano uguali e pari a 1.49618.

4.4.2.3 Fully Informed Particle Swarms (FIPS)

Kennedy e Mendes nel 2002 hanno proposto un'ulteriore versione della PSO, modificando il modo in cui avviene la trasmissione delle informazioni tra le particelle. Il metodo FIPS ha come obiettivo quello di sfruttare l'informazione inutilizzata che si trova nel vicinato di ciascuna particella, intendendo con il termine "vicinato" un determinato intorno di ogni particella. Mentre

¹⁰ Dato che $\phi_1 = \phi_2 = 2.05$, allora nel secondo e terzo addendo si ha $U(0, 2.05)$ e quindi, applicando il moltiplicatore, il vettore di numeri casuali diventa $U(\chi \times 0, \chi \times 2.05) \simeq U(0, 1.49618)$.

nell'algoritmo originale ogni particella è influenzata solamente dalla propria migliore posizione e dalla migliore posizione del vicinato, in questa versione le particelle dipendono dalle migliori posizioni di tutti i loro vicini. L'equazione della PSO è stata quindi modificata ulteriormente nel seguente modo:

$$\begin{cases} V_i^{k+1} = \chi \left(V_i^k + \frac{1}{K_i} \sum_{n=1}^{K_i} U(0, \phi) \otimes (P_{nbr_n^i} - X_i^k) \right), \\ X_i^{k+1} = X_i^k + V_i^{k+1} \end{cases}$$

dove:

- χ è sempre il coefficiente di costrizione;
- ϕ è il coefficiente di accelerazione;
- K_i è il numero di vicini della particella i -esima;
- nbr_n^i rappresenta l' n -esimo vicino della particella i -esima;
- $P_{nbr_n^i}$ per ogni n , rappresenta la migliore posizione raggiunta da ciascuna particella appartenente al vicinato della particella i -esima.

Come si può notare, in questa variante della PSO la velocità della particella i -esima viene aggiornata facendo la media pesata tra le migliori posizioni raggiunte da tutte le particelle appartenenti al vicinato di i . Quindi, mentre nel modello originale l'informazione utilizzata ai fini dello spostamento delle particelle deriva dalla migliore posizione della particella stessa e dalla migliore posizione del vicinato, nella versione FIPS l'informazione risulta essere più completa, in quanto viene fornita da ciascuno dei K_i vicini della particella i -esima.

Analizzando questo modello si può osservare che esso coincide con la versione tradizionale della PSO se K_i viene posto pari a 2. Dai risultati di Kennedy e Mendes, inoltre, risulta che scegliendo dei parametri adeguati per il FIPS è

possibile determinare delle soluzioni migliori e in un numero inferiore di iterazioni rispetto all'algoritmo originale.

4.4.3 Topologia della popolazione

Illustrando l'algoritmo originale della PSO e le sue successive modifiche si è usato varie volte il termine "vicinato", senza però specificare come questo sia strutturato. La scelta di una determinata struttura topologica è cruciale, in quanto definisce il tipo di trasmissione delle informazioni tra le particelle. Generalmente parlando, le topologie possono essere suddivise in due macro-categorie:

- *topologie statiche*: sono caratterizzate dal fatto che la loro struttura rimane immutata all'aumentare delle iterazioni. In particolare, le principali topologie appartenenti a questo gruppo sono le seguenti:
 - topologia *local best*, detta *lbest*, proposta da Kennedy e Eberhart nel 1995. Come si può notare nella Figura 4.2 questa topologia presenta una semplice struttura ad anello, chiamata *ring topology*, in cui ogni particella è connessa e quindi scambia informazioni esclusivamente con i suoi due vicini adiacenti. Sebbene questa struttura del vicinato renda più lungo il processo di ricerca della soluzione ottima, risulta essere meno vulnerabile all'attrazione verso un ottimo locale se confrontata con altre tipologie come la *global best*.

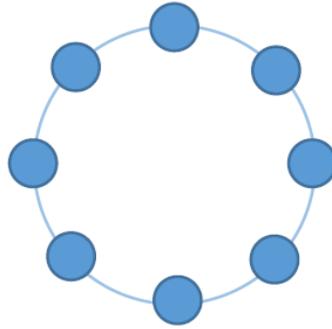


Figura 4.2 - topologia lbest

- topologia *global best*, detta *gbest*. In questa struttura ogni particella è connessa con tutte le altre componenti lo sciame, come si può osservare dalla Figura 4.3. Si tratta quindi di un sistema *fully connected*, dove l'intera popolazione è usata come vicinato di ogni particella e non si generano sub-popolazioni come nel caso precedente. La *gbest* è una struttura che permette una più rapida convergenza dell'algoritmo verso la soluzione ottima, rispetto alla local best, ma può fornire dei risultati peggiori, in quanto vi è una maggiore probabilità che la soluzione individuata dall'algoritmo sia un ottimo locale e non globale. Una maggiore connessione tra le particelle, infatti, se da un lato riduce il tempo computazionale, dall'altro riduce le prestazioni all'aumentare della dimensione del problema in quanto riduce la diversità della popolazione e la capacità di ricerca dell'algoritmo.

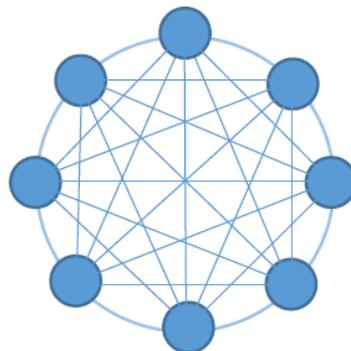


Figura 4.3 - topologia gbest

- topologia di *von Neumann*, che può essere considerata una via di mezzo tra la lbest e la gbest. In questo caso, infatti, la particella i -esima è collegata sia con le particelle adiacenti, sia con quelle che nella struttura circolare del vicinato sono adiacenti alla particella che ha la distanza maggiore da i (Figura 4.4). Kennedy e Mendes nel 2002 hanno dimostrato la superiorità di questa topologia rispetto a quelle illustrate precedentemente.

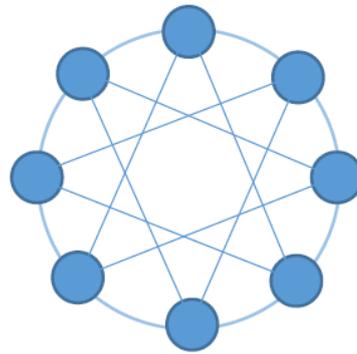


Figura 4.4 - topologia di von Neumann

- *topologie dinamiche*: sono contraddistinte da una struttura del vicinato che si modifica all'aumentare delle iterazioni. Le principali strutture topologiche che si possono trovare in letteratura sono le seguenti:
 - Suganthan (1999), poiché la lbest risulta essere migliore nella fase esplorativa e la gbest converge più velocemente verso la soluzione ottima, propose di cominciare il processo di ricerca con la lbest per poi incrementare gradualmente la numerosità del vicinato fino a raggiungere una topologia di tipo gbest alla fine dell'algoritmo.
 - Jason e Middenford nel 2005 proposero di usare la topologia *hierarchy*, secondo cui vi è una precisa gerarchia tra le particelle dello sciame. Ciascuna particella è soggetta così ad una duplice influenza da parte sia della propria migliore posizione passata, sia

della particella che si posiziona immediatamente sopra di essa nella scala gerarchica per quanto riguarda la qualità del valore della fitness.

- Clerc nel 2006 ha proposto la topologia *Tribes*, dove le caratteristiche, come la dimensione della popolazione, mutano con l'aumentare delle iterazioni a seconda dei risultati ottenuti. In particolare, dopo avere generato diverse sub-popolazioni, si scartano le particelle peggiori che si trovano all'interno delle migliori sub-popolazioni e le si aggiungono alle sub-popolazioni peggiori, incrementando così per queste ultime la probabilità di miglioramento.

4.5 Fireworks Algorithm

Il Fireworks Algorithm (FWA) è un algoritmo basato sulla Swarm Intelligence che è stato recentemente introdotto da Y. Tan e Y. Zhu nel 2010. Esso appartiene alla classe degli algoritmi non-nature inspired, in quanto trae ispirazione dal processo di esplosione dei fuochi d'artificio. Quando un fuoco d'artificio esplose, le scintille si propagano nello spazio circostante e questa esplosione può essere vista come la ricerca nello spazio locale attorno al fuoco d'artificio. L'intera popolazione di fireworks, invece, coopera al fine dell'individuazione della soluzione globale del problema, mettendo così in atto il comportamento sociale di cui si è parlato precedentemente.

Di seguito viene esemplificato il processo algoritmico di ricerca del FWA, effettuando un confronto grafico tra l'esplosione di un fuoco d'artificio (Figura 4.5 a) e il processo di ricerca che avviene in uno spazio delle soluzioni (Figura 4.5 b).

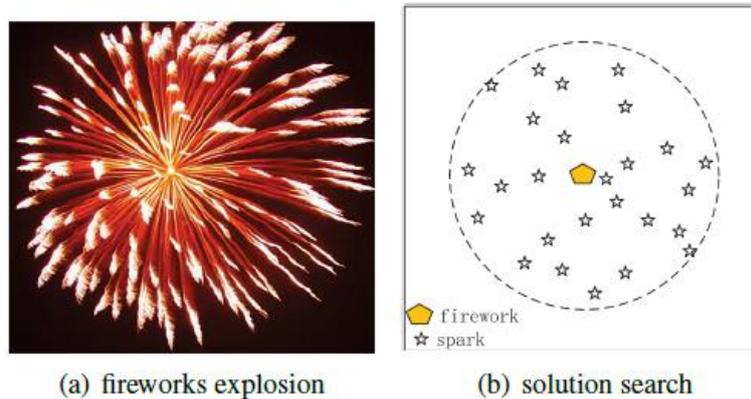


Figura 4.5 - Confronto tra l'esplosione di un fuoco d'artificio e la ricerca dell'ottimo in un problema di ottimizzazione

Il FWA è costituito da quattro componenti chiave, dette *building blocks*, che verranno illustrate nel seguente capitolo: *explosion operator*, *mutation operator*, *mapping rule* e *selection strategy*.

Prima di analizzare nel dettaglio la struttura dell'algoritmo, è necessario evidenziare che il problema di minimizzazione considerato è analogo quello illustrato precedentemente nel capitolo dedicato alla PSO:

$$\min f(\mathbf{x}) \in \mathbb{R}, \quad \mathbf{x}_{min} \leq \mathbf{x} \leq \mathbf{x}_{max},$$

dove:

- $\mathbf{x} = x_1, x_2, \dots, x_n$ è un vettore con n elementi che denota la posizione del singolo fuoco d'artificio nello spazio di ricerca n -dimensionale;
- $f(\mathbf{x})$ è la funzione obiettivo o funzione di fitness;
- \mathbf{x}_{min} e \mathbf{x}_{max} sono vettori che indicano, rispettivamente, il limite inferiore e superiore dello spazio di ricerca in cui si ha l'ammissibilità.

4.6 La struttura del FWA

Il Fireworks Algorithm può essere rappresentato schematicamente dal seguente diagramma di flusso:

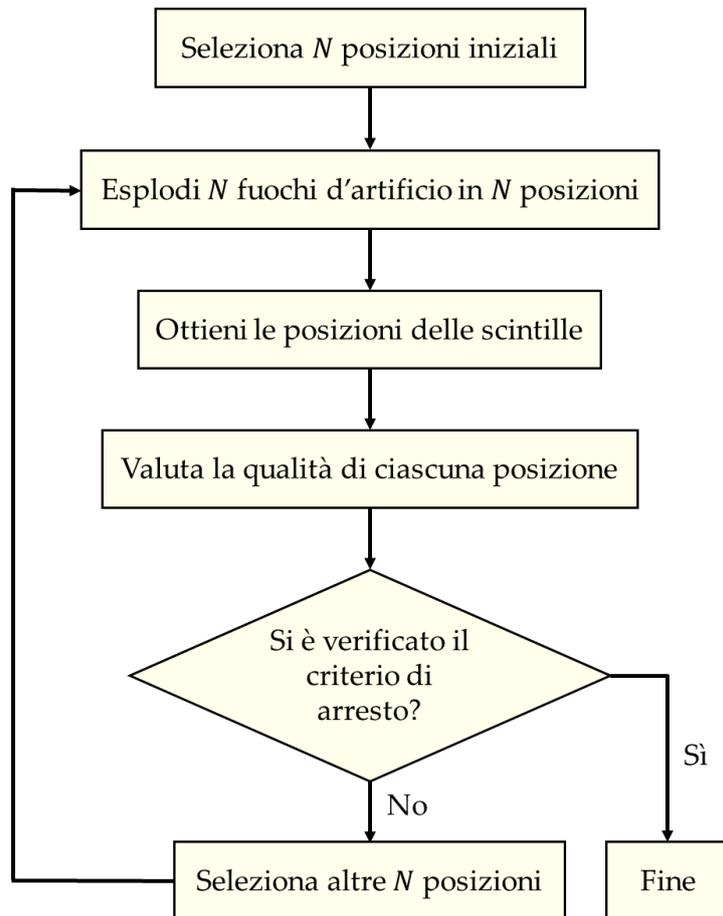


Figura 4.6 - Diagramma di flusso del FWA

Analizzando nella Figura 4.6 la struttura dell'algoritmo, si può notare che lo schema logico delle fasi presenta qualche analogia con quello della PSO, anche se alla base vi sono concetti nettamente diversi.

Il primo passo del FWA consiste nel generare casualmente N posizioni iniziali che sono associate ad altrettanti fuochi d'artificio. I fuochi d'artificio vengono poi fatti esplodere nelle N posizioni selezionate ed è in questa fase che entra in gioco l'*explosion operator*; questa prima componente è a sua volta composta da

tre sotto-componenti, ossia *l'explosion strenght*, *l'explosion amplitude* e il *displacement operation*, che verranno illustrati successivamente.

Tornando al problema di ottimizzazione, le N posizioni iniziali in cui si trovano i fuochi d'artificio sono considerabili come portafogli azionari che vengono generati inizialmente dall'algoritmo. Ciascun portafoglio corrisponde ad una potenziale soluzione ottima del problema e la bontà di una determinata soluzione viene misurata tramite la funzione di fitness che, come è già stato evidenziato, nel nostro caso è l'EVaR che viene sommato alla violazione dei vincoli pesati. Ogni fuoco d'artificio esplodendo produce un determinato numero di scintille, ciascuna associata ad una specifica posizione corrispondente ad un nuovo portafoglio. Dopo l'esplosione, quindi, si avranno sia le posizioni degli N fuochi d'artificio iniziali sia quelle delle scintille prodotte. L'algoritmo così analizza ciascuna soluzione trovata, calcolando la funzione di fitness sia per i fuochi d'artificio che per le scintille prodotte e, se viene soddisfatto un predeterminato criterio di arresto¹¹, il processo si conclude. Se invece tale criterio non viene soddisfatto, a partire dalle posizioni individuate in seguito all'esplosione dei fuochi d'artificio, vengono selezionate, sulla base di due criteri che verranno descritti in seguito, altre N posizioni dalle quali far ripartire un successivo ciclo di esplosioni. In questo modo il processo iterativo si ripete finché non risulta verificato il criterio di arresto.

4.6.1 *Explosion strenght*

Il numero di scintille generate da ciascun fuoco d'artificio viene determinato dalla seguente formulazione relativa all'*explosion strenght*:

¹¹ Può consistere in un soddisfacente valore della funzione di fitness o in un limite massimo di iterazioni.

$$S_i = M_e \cdot \frac{Y_{max} - f(\mathbf{x}_i) + \varepsilon}{\sum_{i=1}^N (Y_{max} - f(\mathbf{x}_i)) + \varepsilon},$$

dove:

- S_i è il numero di scintille generate dall' i -esimo fuoco d'artificio;
- M_e è una costante che controlla il numero totale di scintille generate dagli N fuochi d'artificio;
- $f(\mathbf{x}_i)$ è il valore della funzione obiettivo che assume l' i -esimo fuoco d'artificio;
- $Y_{max} = \max(f(\mathbf{x}_i))$, con $i = 1, \dots, N$, è il massimo valore della funzione obiettivo ottenuto tra tutti gli N fuochi (poiché nel caso in esame si tratta di un problema di minimizzazione, Y_{max} è il peggior valore che può assumere la funzione);
- ε è una costante trascurabile, usata per evitare che si annulli il valore delle quantità al numeratore e al denominatore.

Per evitare che il numero di scintille prodotte sia troppo elevato, situazione che rallenterebbe il processo di ricerca dell'ottimo, oppure che sia troppo basso, situazione che ridurrebbe la capacità esplorativa dell'algoritmo, è inoltre necessario definire i seguenti vincoli per S_i :

$$\hat{S}_i = \begin{cases} \text{round}(a \cdot M_e) & \text{se } S_i < aM_e \\ \text{round}(b \cdot M_e) & \text{se } S_i > bM_e, \\ \text{round}(S_i) & \text{altrimenti} \end{cases}$$

dove:

- \hat{S}_i è la limitazione del numero delle scintille;
- a e b sono costanti, con $0 < a < b < 1$;
- $\text{round}()$ è una funzione che permette di arrotondare l'argomento all'intero più vicino.

Osservando la formulazione dell'explosion strenght si può notare che l'algoritmo genera per ciascun fuoco d'artificio un numero di scintille tanto più elevato quanto migliore è la relativa funzione di fitness. Infatti, in ogni iterazione del processo, il denominatore risulta essere costante mentre il numeratore varia a seconda del valore della funzione calcolata in ciascun firework; tanto più questo valore è peggiore, ossia elevato, quanto più il numeratore decresce, e di conseguenza anche il numero di scintille prodotte dal firework in questione si riduce. In questo modo l'algoritmo approfondisce maggiormente la ricerca della soluzione ottima in zone dello spazio più promettenti, dove più probabilmente si troverà la soluzione desiderata.

4.6.2 Explosion amplitude

L'algoritmo, dopo aver calcolato il numero di scintille, calcola l'ampiezza di ciascuna esplosione:

$$A_i = \hat{A} \cdot \frac{f(\mathbf{x}_i) - Y_{min} + \varepsilon}{\sum_{i=1}^N (f(\mathbf{x}_i) - Y_{min}) + \varepsilon},$$

dove:

- A_i è l'ampiezza dell'esplosione dell' i -esimo firework;
- \hat{A} è una costante che indica l'ampiezza totale delle esplosioni degli N fuochi d'artificio;
- $Y_{min} = \min(f(\mathbf{x}_i))$, con $i = 1, \dots, N$, è il minimo valore della funzione obiettivo ottenuto tra tutti gli N fuochi d'artificio, corrispondente al miglior valore della funzione di fitness;
- $f(\mathbf{x}_i)$ e ε sono già stati illustrati precedentemente.

Si può notare che l'ampiezza dell'esplosione, al contrario del numero di scintille, è tanto minore quanto migliore è la funzione di fitness di un determinato fuoco d'artificio. Ciò è coerente con quanto evidenziato nel capitolo dedicato alle metaeuristiche, dove si è parlato del bilanciamento tra esplorazione e sfruttamento nel processo algoritmico di ricerca. In questi termini si può comprendere come una buona posizione del singolo firework spinga l'algoritmo ad un maggiore sfruttamento dello spazio ammissibile, approfondendo la ricerca localmente. Al contrario, se la funzione di fitness rileva una pessima posizione del firework, la ricerca viene spinta in zone dello spazio più lontane, continuando così il processo esplorativo.

Si può così comprendere come la posizione di ciascun firework determini il numero di scintille da generare e l'ampiezza dell'esplosione, suddividendo i fuochi d'artificio in due categorie:

- *fireworks di buona qualità*: essi generano buone esplosioni di ampiezza ridotta con un elevato numero di scintille (Figura 4.7 a);
- *fireworks di cattiva qualità*: questi generano invece esplosioni di ampiezza elevata con una popolazione ridotta di scintille (Figura 4.7 b).

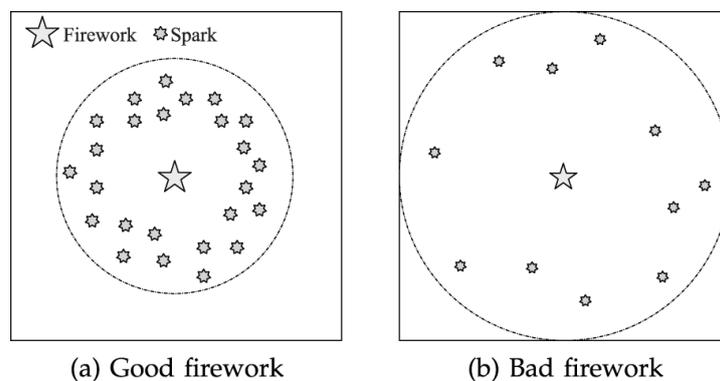


Figura 4.7 - Tipologie di fireworks

Queste due tipologie di fireworks vengono dunque incontro ad una duplice necessità dell'algoritmo, come già è stato evidenziato poco sopra. In primo luogo quelli di cattiva qualità esplorano lo spazio di ricerca, evitando di far convergere l'algoritmo troppo velocemente verso un possibile ottimo e cercando di trovare zone più promettenti nello spazio ammissibile. Quelli di qualità migliore, invece, poiché presumibilmente dovrebbero essere più vicini all'ottimo del problema, approfondiscono la ricerca in una determinata zona dello spazio che risulta essere più promettente.

4.6.3 *Displacement operation*

L'ultima fase dell'explosion operator va sotto il nome di displacement operation, che indica un'operazione di spostamento. Questa fase è estremamente importante, in quanto la posizione di ciascun firework è associata ad un vettore n -dimensionale e quindi il problema da risolvere è multidimensionale. Nel caso specifico, quando si vogliono calcolare le posizioni delle scintille, ai fini di ottenere una maggiore aleatorietà nelle esplosioni, non tutte le dimensioni e quindi le componenti del vettore associato all' i -esimo firework vengono coinvolte nell'operazione di spostamento. Questo spostamento viene indicato con h e, considerando un generico problema n -dimensionale, viene applicato solo a $z < n$ dimensioni scelte casualmente secondo il seguente criterio:

$$z = \text{round}(n \cdot \text{rand}(0, 1)),$$

dove:

- n è il numero delle dimensioni del problema, ossia degli elementi che compongono il vettore che identifica la posizione di ciascun firework;

- $rand(0,1)$ è una variabile casuale uniforme distribuita nell'intervallo $[0,1]$.

Successivamente vengono scelte casualmente le z dimensioni tra le n esistenti. Da un punto di vista geometrico solo per le z dimensioni estratte avverrà uno spostamento all'interno dello spazio di ricerca e quindi le esplosioni coinvolgeranno solo alcune dimensioni.

Lo spostamento h , per ognuna delle z dimensioni, viene poi calcolato tenendo conto dell'ampiezza A_i relativa alla posizione del fuoco x_i :

$$h = A_i \cdot rand(-1, 1),$$

dove $rand(-1,1)$ è una variabile casuale uniforme distribuita nell'intervallo $[-1,1]$. In questo modo, entro i limiti dell'ampiezza A_i relativa all'esplosione dell' i -esimo fuoco d'artificio, ciascuna delle z dimensioni di ogni fuoco subisce uno spostamento casuale che conferisce ulteriore aleatorietà al processo.

Se l'ampiezza delle esplosioni non cambiasse tra una dimensione e l'altra, si avrebbero ad esempio nel caso bidimensionale esplosioni che formano circonferenze perfette nello spazio, in quanto il raggio sarebbe costante. Grazie al fattore $rand(-1,1)$, perciò, ciascuna dimensione subisce uno spostamento casuale entro i limiti dell'ampiezza determinata, incrementando così l'esplorazione dello spazio di ricerca.

La nuova posizione del fuoco d'artificio i -esimo con riferimento alla j -esima dimensione tra le z che sono state selezionate è quindi calcolata come:

$$x_i^{j+1} = x_i^j + h,$$

dove x_i^j è la posizione corrente dell' i -esimo firework con riferimento alla j -esima posizione.

4.6.4 Mutation operator

Il secondo building block del Fireworks Algorithm va sotto il nome di mutation operator. Esso garantisce ulteriormente la diversità dello sciame attraverso delle esplosioni che generano nuove scintille a partire dagli N fireworks iniziali. In questo modo la ricerca locale promossa dall'explosion operator viene affiancata da un'ulteriore tipologia di esplosioni finalizzate a migliorare la ricerca globale della soluzione ottima. Le nuove scintille vengono generate dall'algoritmo grazie ad un operatore chiamato *Gaussian operator*; mediante quest'ultimo si ottengono degli spostamenti casuali di ciascun fuoco d'artificio, moltiplicando un valore casuale estratto da una distribuzione di probabilità Normale con la posizione del firework relativamente a $z < n$ dimensioni¹². In questo caso il meccanismo di spostamento è più semplice, dato che l'ampiezza dell'esplosione e il numero di scintille prodotte non sono influenzate dalla posizione del fuoco d'artificio. Il numero di scintille generate per ogni fuoco d'artificio dall'esplosione Gaussiana è infatti posto pari alla costante M_g , mentre l'ampiezza di ciascuna esplosione è casuale.

Definita x_i^j la posizione nello spazio dell' i -esimo fuoco d'artificio, relativamente alla j -esima dimensione tra le z che sono state selezionate, lo spostamento viene calcolato nel seguente modo:

$$x_i^{j+1} = x_i^j \cdot g,$$

¹² Le z dimensioni sono selezionate casualmente secondo la formulazione già vista in precedenza, ossia $z = \text{round}(n \cdot \text{rand}(0, 1))$.

dove g è un numero casuale estratto da una distribuzione Gaussiana, o Normale, con media 1 e varianza 1:

$$g = \text{Gaussian}(1, 1).$$

Questa operazione viene ripetuta M_g volte per ciascuno degli N fuochi d'artificio e ogni posizione viene aggiornata con riferimento ad ognuna delle z dimensioni selezionate casualmente.

4.6.5 *Mapping rule*

La terza componente dell'algoritmo consiste in una rimappatura delle posizioni delle scintille che cadono al di fuori dello spazio ammissibile di ricerca. Questa fase serve essenzialmente a gestire le violazioni dei vincoli di bounds. Ai fini del presente lavoro, però, la mapping rule risulta superflua, in quanto la violazione di tutti i vincoli del problema di ottimizzazione, compresi i vincoli di bounds, viene gestita mediante il metodo delle penalità esatte, che verrà illustrato nel successivo capitolo.

4.6.6 *Selection strategy*

La quarta componente del Fireworks Algorithm prende il nome di selection strategy e serve per selezionare le posizioni delle scintille che diventeranno le nuove posizioni dei fuochi d'artificio nell'iterazione successiva. Quest'ultima fase avviene se non viene soddisfatto il prestabilito criterio di arresto. In caso di non soddisfacimento del criterio di arresto l'algoritmo, per selezionare le nuove posizioni dove fare esplodere i nuovi N fuochi d'artificio, tra le varie si utilizza

una strategia che va sotto il nome di *distance-based strategy*. Innanzitutto, data la posizione attuale, la scintilla con la migliore posizione, \mathbf{x}^* , viene sempre utilizzata nell'iterazione successiva e viene quindi fatta nuovamente esplodere secondo i criteri visti in precedenza. Quindi, poiché in ogni iterazione vengono fatti esplodere N fuochi d'artificio, rimangono $(N - 1)$ posizioni da selezionare attraverso la *distance-based strategy*. Questa strategia consiste nel selezionare le nuove posizioni sulla base della distanza tra ciascuna scintilla e tutte le altre. In generale, la distanza tra la posizione \mathbf{x}_i e le altre posizioni è calcolata nel seguente modo:

$$R(\mathbf{x}_i) = \sum_{j=1}^N d(\mathbf{x}_i, \mathbf{x}_j) = \sum_{j=1}^N \|\mathbf{x}_i, \mathbf{x}_j\|,$$

dove:

- $R(\mathbf{x}_i)$ è la somma delle distanze tra \mathbf{x}_i e le altre scintille;
- $d(\mathbf{x}_i, \mathbf{x}_j)$ è la distanza euclidea¹³ tra la scintille \mathbf{x}_i e \mathbf{x}_j , che può essere scritta anche con la notazione relativa alla norma euclidea¹⁴ $\|\mathbf{x}_i, \mathbf{x}_j\|$.

La distanza calcolata viene poi utilizzata per calcolare la probabilità che la posizione \mathbf{x}_i venga selezionata dall'algoritmo come nuova posizione dove far esplodere un fuoco d'artificio nella successiva iterazione. Tale probabilità viene calcolata nella maniera seguente:

$$p(\mathbf{x}_i) = \frac{R(\mathbf{x}_i)}{\sum_{j \in K} R(\mathbf{x}_j)},$$

¹³ In matematica, la distanza euclidea è la distanza tra due punti, ossia la misura del segmento avente per estremi i due punti. Il concetto di distanza si può genericamente estendere ad uno spazio n -dimensionale.

¹⁴ Una *norma* è una funzione che assegna ad ogni vettore di uno spazio vettoriale, tranne lo zero, una lunghezza positiva. La norma di indice 2, detta euclidea, viene definita come $\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$.

dove l'indice $j \in K$ indica che la posizione j -esima appartiene all'insieme K comprensivo di tutte le posizioni generate dall'algoritmo, relative sia alle scintille che ai fuochi.

In questo modo la probabilità di selezionare una certa posizione per l'iterazione successiva dipende dalla somma delle distanze euclidee tra la posizione associata a quella scintilla e le posizioni di tutte le altre scintille dello sciame. Quindi, poiché il denominatore della frazione rimane costante, le scintille situate in posizioni più distanti rispetto alle altre hanno maggiore probabilità di essere scelte nella successiva iterazione e di dar vita ad una successiva esplosione. Al contrario, le scintille che si trovano in posizioni più vicine agli altri elementi dello sciame hanno una probabilità minore di diventare fuochi d'artificio nella successiva iterazione, e quindi di dare origine ad una successiva generazione di scintille. Con quest'ultima fase del processo algoritmico, dunque, si vuole provocare esplosioni di fuochi d'artificio nelle zone dello spazio ammissibile che fino a quel momento sono state meno esplorate; in questo modo viene garantita la diversità e l'eterogeneità della popolazione e si ha un maggiore capacità esplorativa dell'algoritmo.

4.6.7 Pseudo-codice FWA

Si riporta di seguito uno pseudo-codice del Fireworks Algorithm:

- 1: Inizializza N posizioni e calcola la funzione di fitness $f(x_i)$ per ogni firework;
- 2: **fino a che** le condizioni di arresto non sono raggiunte
- 3: **per** ogni posizione x_i , con $i = 1, \dots, N$
- 4: Calcola la funzione di fitness $f(x_i)$;
- 5: Calcola il numero di scintille S_i ;

- 6: Calcola l'ampiezza dell'esplosione A_i ;
- 7: Seleziona casualmente $z < n$ dimensioni relative alla posizione x_i ;
%displacement operation
- 8: **per** ogni dimensione $j \in z$ di ogni x_i
- 9: Calcola lo spostamento $h = A_i \cdot rand(-1, 1)$;
- 10: Calcola la nuova posizione $x_i^{j+1} = x_i^j + h$;
- 11: **fine ciclo per**
%mutation operator
- 12: **per** $k = 1, \dots, M_g$
- 13: Seleziona casualmente $z < n$ dimensioni relative alla posizione x_i ;
- 14: **per** ogni dimensione $j \in z$ di ogni x_i
- 15: Calcola il coefficiente dell'esplosione Gaussiana: $g = Gaussian(1, 1)$;
- 16: Calcola le scintille dell'esplosione Gaussiana: $x_i^{j+1} = x_i^j \cdot g$;
- 17: **fine ciclo per**
- 18: **fine ciclo per**
- 19: Seleziona la posizione della scintilla con il più basso valore della funzione di fitness e le altre posizioni secondo la probabilità: $p(x_i) = \frac{R(x_i)}{\sum_{j \in K} R(x_j)}$.
- 20: **fine del ciclo fino a che**

4.7 Caratteristiche del Fireworks Algorithm

Alla luce di quanto illustrato nelle precedenti pagine, è possibile individuare alcune caratteristiche che contraddistinguono il Fireworks Algorithm:

- *esplosività*: i fuochi d'artificio esplodono generando un certo numero di scintille. La selection strategy, poi, individua N scintille che nell'iterazione successiva diventano le posizioni dei nuovi fireworks;
- *istantaneità*: le scintille hanno "vita breve", ossia quelle che non vengono selezionate con la selection strategy svaniscono;
- *semplicità*: ogni firework possiede informazioni solo in merito alla propria posizione;
- *località*: ciascun firework genera scintille solo entro i limiti della propria ampiezza;
- *proprietà di emergenza della complessità*: questa proprietà è legata al concetto di Swarm Intelligence, infatti la popolazione di fireworks nel suo insieme dà vita ad un comportamento estremamente più complesso rispetto a quello dei singoli fuochi d'artificio;
- *parallelismo*: i fireworks sono indipendenti e non c'è un meccanismo centrale che regola l'intero processo esplosivo perché gli individui dello sciame interagiscono tra di loro e si scambiano informazioni, secondo i principi propri della Swarm Intelligence;
- *diversità*: questa è una delle caratteristiche più importanti, in quanto consente all'algoritmo di convergere all'ottimo globale senza rimanere "intrappolato" in un punto di ottimo locale. La diversità può essere individuata nel numero di scintille generate e nell'ampiezza delle esplosioni, che cambia a seconda della posizione del singolo fuoco d'artificio. La diversità viene poi garantita dal displacement operation e dal mutation operator, oltre che dalla selection strategy dove, come si è

appena visto, le scintille più distanti dalle altre hanno una maggiore di probabilità di essere scelte, al fine di preservare la diversità della popolazione;

- *robustezza*: quest'ultima caratteristica è anch'essa importante, in quanto evidenzia il fatto che il FWA non è influenzato dal comportamento dei peggiori fireworks, in quanto i fuochi d'artificio sono tra loro indipendenti.

4.8 Confronto tra PSO e FWA

Nel seguente paragrafo effettueremo alcuni confronti tra la Particle Swarm Optimization nella versione con il peso d'inerzia e il Fireworks Algorithm, algoritmi che verranno implementati nel successivo capitolo per risolvere il problema di selezione di portafoglio.

Entrambi gli algoritmi sono ispirati alla Swarm Intelligence, tuttavia la PSO è una metaeuristica bio-ispirata, mentre il FWA è una metaeuristica non bio-ispirata. Il processo algoritmico nei due differenti casi risulta essere in parte simile e in parte differente, come viene di seguito evidenziato:

1. Nella fase iniziale Il FWA inizializza N posizioni corrispondenti ai fuochi d'artificio iniziali, mentre la PSO inizializza M particelle.
2. Nella seconda fase entrambi gli algoritmi calcolano la funzione di fitness per ogni posizione, intendendo così la posizione del fuoco d'artificio per il FWA e la posizione della particella per la PSO.
3. Nella terza fase vi è un maggior numero di differenze. La PSO confronta il valore della funzione di fitness associata alla posizione della particella corrente con la migliore posizione passata: se la funzione di fitness corrente è minore allora quest'ultima verrà considerata la nuova migliore

posizione della particella. Viene inoltre valutata la migliore posizione del vicinato e infine vengono aggiornate la velocità e la posizione delle particelle. Nel caso del FWA, invece, la posizione delle scintille viene determinata mediante gli operatori displacement operation e mutation operator. La miglior funzione di fitness associata alla migliore posizione viene memorizzata per l'iterazione successiva. Infine, la probabilità delle scintille di essere selezionate come individui della popolazione successiva dipende dalla somma delle distanze euclidee tra ciascuna di esse e tutte le altre scintille dello sciame.

Le strategie di selection strategy sono perciò diverse nei due algoritmi e questa diversità è anche data dalla differenza delle teorie a cui esse si ispirano. Il FWA si ispira all'*immune concentration theory*: ogni scintilla viene trattata come un anticorpo di un sistema immunitario. Una scintilla (anticorpo) che ha molte scintille (anticorpi) simili, ossia vicine, ha una probabilità minore di essere scelta rispetto a quelle che hanno invece una distanza maggiore dalle altre. Questo meccanismo preserva la diversità della popolazione nel FWA che non è invece presente nella PSO.

Per quanto riguarda invece i parametri, riportiamo di seguito l'elenco dei parametri che è necessario settare inizialmente nei due algoritmi:

- Nella PSO le costanti da determinare sono:
 - il numero di particelle M ;
 - il numero massimo di iterazioni K ;
 - i due coefficienti di accelerazione ϕ_1 e ϕ_2 ;
 - ω_{min} e ω_{max} per il calcolo del peso d'inerzia che decresce con il numero delle iterazioni;
 - V_{max} , ossia la velocità massima.

- Nel FWA le costanti da settare sono invece:
 - il numero di fuochi d'artificio N ;
 - il numero massimo di iterazioni K ;
 - l'ampiezza totale delle esplosioni \hat{A} ;
 - M_e , il numero totale di scintille generate dagli N fuochi d'artificio con l'explosion strenght;
 - i parametri a e b che controllano il numero minimo e massimo di scintille generate con l'explosion strenght;
 - M_g , il numero di scintille prodotte dall'esplosione Gaussiana per ogni fuoco d'artificio.

I due parametri comuni a tutti e due gli algoritmi sono la dimensione della popolazione iniziale e il massimo numero di iterazioni consentite.

Nelle figure 4.9 e 4.10 si riporta infine un'esemplificazione grafica del movimento delle particelle e delle scintille, per cogliere meglio le differenze tra i due diversi processi.

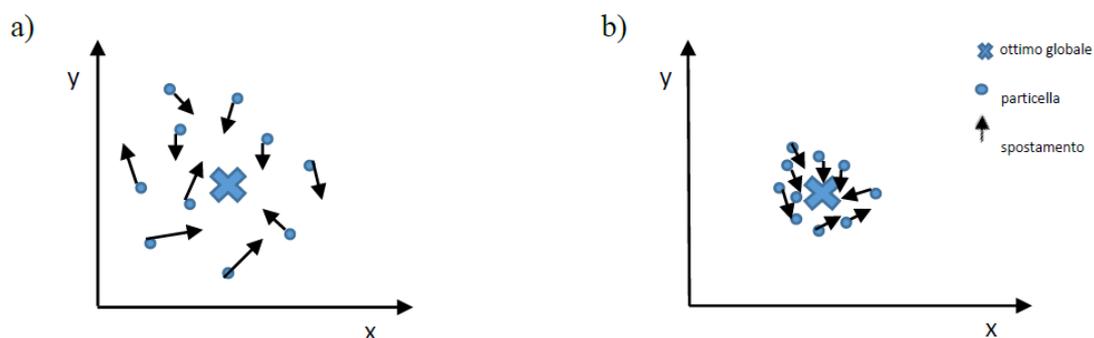


Figura 4.9 - Esempio bidimensionale del movimento delle particelle: in a) si osserva il movimento iniziale delle particelle che in b) convergono all'ottimo globale.

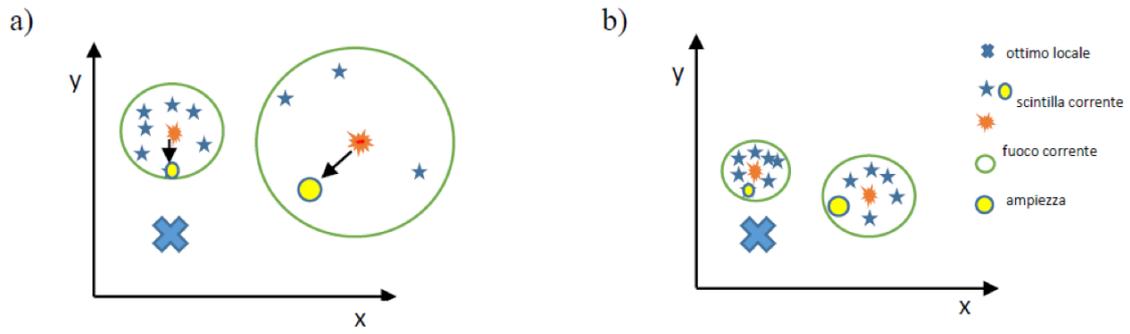


Figura 4.10 - Esempio bidimensionale del movimento delle scintille verso l'ottimo. In a) avvengono le prime esplosioni e successivamente vengono selezionate le scintille che saranno le posizioni dei nuovi fuochi d'artificio nell'iterazione successiva, come si può vedere in b).

CAPITOLO 5

APPLICAZIONE

5.1 Riformulazione del problema di selezione di portafoglio

La PSO e il FWA, illustrate nel precedente capitolo, sono due metaeuristiche in grado di risolvere problemi di ottimizzazione non vincolata. Poiché il problema di selezione di portafoglio proposto nella presente tesi è di natura vincolata¹, al fine di usare come tecniche risolutive i due algoritmi riportati sopra, l'approccio qui utilizzato consiste nel riformulare il problema vincolato in uno non vincolato. Per far ciò, tra le varie tecniche proposte in letteratura, si è scelto di utilizzare il metodo delle penalità esatte (exact penalty method). Tale metodo prevede l'introduzione di una funzione di penalità, costruita aggiungendo alla funzione obiettivo un termine che penalizza la violazione dei vincoli. In sostanza, alla funzione obiettivo viene aggiunta la somma di tutte le violazioni dei vincoli moltiplicate per uno specifico parametro $\frac{1}{\varepsilon} > 0$. In questo modo, oltre a minimizzare la funzione in questione vengono minimizzate anche tutte le violazioni dei vincoli.

¹ Si veda il paragrafo 3.3.

5.1.1 Metodo delle penalità esatte

Formalmente, il metodo delle penalità esatte consiste nel riscrivere un generico problema di ottimizzazione vincolata del seguente tipo (Luenberger, 1984²):

$$\begin{aligned} & \min f(x) \\ \text{s. t. } & \begin{cases} h(x) = 0 \\ g(x) \leq 0 \end{cases} \end{aligned}$$

dove:

- $f(x)$ è una funzione continua in \mathbb{R}^n ;
- $h(x)$ rappresenta m vincoli espressi sotto forma di uguaglianza;
- $g(x)$ rappresenta p vincoli definiti sotto forma di disuguaglianza,

come un problema di ottimizzazione non vincolata formulato nella seguente maniera:

$$\min f(x) + \frac{1}{\varepsilon} P(x), \quad \text{con } P(x) \geq 0 \quad \forall x \in \mathbb{R}^n,$$

dove:

- $\frac{1}{\varepsilon}$ è una costante positiva chiamata fattore di penalità;
- $P(x)$ è la funzione di penalità che è continua in \mathbb{R}^n .

Il ruolo del fattore di penalità è quello di aumentare il valore numerico di ogni singola violazione. È indispensabile scegliere correttamente il valore di questo parametro, per avere la corrispondenza tra la soluzione del problema vincolato e di quello riformulato.

² David G. Luenberger, *Linear and Nonlinear Programming*, Second Edition, Addison-Wesley, Inc., Reading, Massachusetts, 1984.

In particolare, se si formula la funzione di penalità nel seguente modo:

$$P(x) = \sum_{i=1}^m |h_i(x)| + \sum_{j=1}^p \max\{0, g_j(x)\},$$

allora se un generico punto x^* soddisfa le condizioni sufficienti del secondo ordine³ per essere un punto di minimo locale del problema vincolato, tale punto sarà un minimo locale anche nel problema di ottimizzazione non vincolata se si ha soddisfatta la seguente relazione:

$$\frac{1}{\varepsilon} > \max\{|\omega_i|, \mu_j\}, \text{ con } i = 1, \dots, m \text{ e } j = 1, \dots, p,$$

dove ω e μ sono due vettori composti dai moltiplicatori di Lagrange⁴ relativi, rispettivamente, agli m vincoli compresi in $h(x) = 0$ ed ai p vincoli compresi nel vettore $g(x) \leq 0$.

Questo risultato teorico, come si può notare, fornisce le condizioni sufficienti affinché si possa avere coincidenza tra le soluzioni del problema vincolato e di quello non vincolato, ma non fornisce alcuna indicazione in merito alla determinazione del parametro ε .

Tornando al nostro problema di selezione di portafoglio, è quindi necessario riformulare i vincoli enunciati nel paragrafo 3.3 nella maniera seguente:

³ Le condizioni sufficienti del secondo ordine per problemi di ottimizzazione vincolata affermano che, se la matrice hessiana delle derivate seconde della funzione obiettivo risulta essere definita positiva (o negativa), allora il punto stazionario individuato è un punto di minimo (o di massimo) locale della funzione.

⁴ I moltiplicatori di Lagrange sono dei parametri che vengono utilizzati nel metodo dei moltiplicatori di Lagrange per risolvere problemi di ottimizzazione vincolata con vincoli espressi sottoforma di equazioni. Tali parametri, indicati con la lettera λ , permettono di riscrivere la funzione obiettivo in una funzione lagrangiana dove vi è la presenza anche dei vincoli che vengono moltiplicati per queste nuove variabili scalari.

$$\begin{aligned}
\sum_{i=1}^N x_i = 1 &\rightarrow \left| \sum_{i=1}^N x_i - 1 \right|, \\
\hat{R}_p \geq \pi &\rightarrow \max\{0, \pi - \hat{R}_p\}, \\
K_l \leq \sum_{i=1}^N z_i &\rightarrow \max\left\{0, K_l - \sum_{i=1}^N z_i\right\}, \\
\sum_{i=1}^N z_i \leq K_u &\rightarrow \max\left\{0, \sum_{i=1}^N z_i - K_u\right\}, \\
z_i l_i \leq x_i &\rightarrow \sum_{i=1}^N \max\{0, z_i l_i - x_i\}, \\
x_i \leq z_i u_i &\rightarrow \sum_{i=1}^N \max\{0, x_i - z_i u_i\}, \\
z_i(z_i - 1) = 0 &\rightarrow \sum_{i=1}^N |z_i(z_i - 1)|.
\end{aligned}$$

Ora è possibile riformulare il problema di selezione iniziale in forma non vincolata nel seguente modo:

$$\min_{X, Z, \varepsilon} P(X, Z, \varepsilon),$$

dove:

$$\begin{aligned}
P(X, Z, \varepsilon) &= EVaR_{1-\alpha}(R_p) \\
&+ \frac{1}{\varepsilon} \left[\left| \sum_{i=1}^N x_i - 1 \right| + \max\{0, \pi - \hat{R}_p\} + \max\left\{0, K_l - \sum_{i=1}^N z_i\right\} \right. \\
&+ \max\left\{0, \sum_{i=1}^N z_i - K_u\right\} \\
&\left. + \sum_{i=1}^N \max\{0, z_i l_i - x_i\} + \sum_{i=1}^N \max\{0, x_i - z_i u_i\} + \sum_{i=1}^N |z_i(z_i - 1)| \right].
\end{aligned}$$

Questo modello è dunque la riformulazione non vincolata del problema di minimizzazione proposto in questo lavoro. Si può notare che la somma algebrica all'interno delle parentesi quadre è nulla solo nel caso in cui vengono rispettati tutti i vincoli. Al contrario, basta la violazione di un vincolo per rendere tale somma una quantità positiva⁵ che, amplificata dal fattore moltiplicativo $\frac{1}{\varepsilon}$, si sommerebbe alla misura di rischio. Ad esempio, il vincolo legato al rendimento medio minimo del portafoglio, se violato, porterebbe ad avere un rendimento medio effettivo (\hat{R}_p) minore rispetto al rendimento target (π); di conseguenza la quantità ($\max\{0, \pi - \hat{R}_p\}$) sarebbe positiva e pari alla differenza ($\pi - \hat{R}_p$). Quest'ultima quantità, moltiplicata per $\frac{1}{\varepsilon}$, si andrebbe a sommare all'Entropic Value-at-Risk generando così un valore maggiore della funzione di fitness riformulata. Gli algoritmi, perciò, cercano di evitare le soluzioni che non rispettano i vincoli, in quanto la violazione di questi ultimi comporta un aumento notevole del valore della fitness, che dev'essere minimizzata.

5.2 Due casi a confronto

Nel seguente paragrafo e in quelli a seguire verrà concretamente affrontato il problema di ottimizzazione di cui si è ampiamente parlato fino ad ora. Le metaeuristiche usate a fini risolutivi sono la PSO e il FWA⁶, che sono state implementate in MATLAB R2019b.

Si è scelto di testare la bontà del modello costruito nei precedenti capitoli sui dati reali relativi a due mercati finanziari: quello italiano e quello svedese. In

⁵ L'operatore *valore assoluto* e la funzione *massimo* rendono inevitabilmente tutte le quantità positive, o al massimo nulle.

⁶ Per i codici Matlab si rimanda all'Appendice A per la PSO e all'Appendice B per il FWA.

particolare, per il primo si è preso in considerazione l'indice azionario FTSE MIB, che è il principale benchmark dei mercati azionari italiani in quanto comprende le 40 società italiane quotate con maggiore capitalizzazione, liquidità e flottante che rappresentano circa l'80% della capitalizzazione del mercato italiano. Per il mercato svedese si è scelto invece l'indice OMX Stockholm 30, il principale indice azionario della Borsa di Stoccolma che racchiude i 30 titoli svedesi più scambiati e di conseguenza a maggiore liquidità.

La scelta di questi due mercati è stata indotta dal fatto che si tratta di due Paesi in situazioni economiche completamente differenti. L'Italia negli ultimi anni ha attraversato un periodo economico caratterizzato da una profonda incertezza, alternando periodi di recessione a periodi di stagnazione. La crescita del Pil è sostanzialmente nulla e il debito pubblico ha raggiunto un ammontare record. L'economia italiana, come sottolineato dalla Nota mensile prodotta dall'Istat nello scorso giugno, è inoltre caratterizzata da un proseguimento della fase di debolezza dei ritmi produttivi che però è associata a miglioramenti sul mercato del lavoro e del potere di acquisto delle famiglie. Anche l'indice di fiducia dei consumatori e quello relativo alla fiducia delle imprese hanno mostrato un sensibile peggioramento nei recenti periodi, causato appunto dalla congiuntura economica italiana.

La Svezia invece, anche se negli ultimi anni ha subito un rallentamento della crescita economica, è una delle maggiori economie a livello sia europeo che internazionale. È caratterizzata da un'alta qualità della vita, da infrastrutture solide e da un ottimo sistema di assistenza sanitaria e istruzione. La Svezia ha inoltre saputo reagire efficacemente contro la crisi finanziaria degli anni '90, che ha visto il freno della crescita del Pil, un aumento incontrollato della disoccupazione e dei tassi di interesse. In tali circostanze, per far ripartire la crescita economica, il governo svedese ha scelto la privatizzazione dei taxi, dell'elettricità, delle telecomunicazioni e delle ferrovie e una serie di riforme

hanno poi reso più semplice la concessione di licenze per la costituzione di nuove società e hanno contribuito a far uscire dal mercato le aziende inefficienti. Il governo svedese ha inoltre promosso l'idea secondo cui tagliare le aliquote d'imposta societarie può contribuire a stimolare l'imprenditorialità e questa serie di iniziative hanno contribuito ad una ripresa di tale economia, sino alla creazione di una potenza economica internazionale.

Si è perciò scelto, ai fini della specificazione del nostro modello di portafoglio, di usare i dati relativi ai prezzi azionari delle principali società di questi due Paesi, che versano in due situazioni economiche nettamente differenti. In particolare si è scelto di usare le serie storiche dei prezzi di chiusura giornalieri⁷ relativi all'arco temporale che va dal 01/01/2015 al 30/06/2019.

Di seguito sono riportati i nomi delle società emittenti e i relativi tickers⁸, con riferimento ai titoli che compongono i due indici azionari. In particolare, tra i titoli componenti il FTSE MIB, non sono stati tenuti in considerazione Ferrari (RACE.MI), Italgas (IG.MI), Nexi (NEXI.MI), Pirelli & C. (PIRC.MI) e Poste Italiane (PST.MI), in quanto alcune di queste società si sono quotate in un periodo successivo al 01/01/2015 e altre sono state oggetto di operazioni straordinarie che hanno provocato il delisting⁹ e una successiva seconda quotazione in Borsa posteriore all'inizio del periodo campionario di riferimento. Le società usate ai fini della specificazione del modello per l'indice italiano sono le seguenti: A2A (A2A.MI), Amplifon (AMP.MI), Atlantia (ATL.MI), Azimut Holding (AZM.MI), Banco BPM (BAMI.MI), BPER Banca (BPE.MI), Buzzi Unicem (BZU.MI), Campari (CPR.MI), CNH Industrial (CNHI.MI), DiaSorin (DIA.MI), Enel (ENEL.MI), Eni (ENI.MI), Exor (EXO.MI), Fiat Chrysler Automobiles (FCA.MI), FincoBank (FBK.MI), Generali (G.MI),

⁷ I dati sono stati scaricati con la piattaforma Bloomberg.

⁸ Il ticker, o sigla di negoziazione, è un codice identificativo utilizzato per la negoziazione in Borsa.

⁹ Il delisting è la revoca della società dalle negoziazioni di Borsa.

Hera (HER.MI), Intesa Sanpaolo (ISP.MI), Juventus Football Club (JUVE.MI), Leonardo (LDO.MI), Mediobanca (MB.MI), Moncler (MONC.MI), Prysmian (PRY.MI), Recordati (REC.MI), Saipem (SPM.MI), Salvatore Ferragamo (SFER.MI), Snam (SRG.MI), STMicroelectronics (STM.MI), Telecom Italia (TIT.MI), Tenaris (TEN.MI), Terna - Rete Elettrica Nazionale (TRN.MI), UBI Banca (UBI.MI), UniCredit (UCG.MI), Unipol (UNI.MI), UnipolSai Assicurazioni (US.MI).

Per quanto riguarda l'OMX Stockholm 30, invece, sono stati presi in considerazione tutti i titoli, ad eccezione di ESSITY AKTIEBOLAG-B (ESSITYB SS Equity), società attiva nel campo dell'igiene e della salute personale che si è quotata alla Borsa di Stoccolma il 15/06/2017, quindi in tempi più recenti. Le altre società componenti l'indice sono le seguenti: Astrazeneca Plc (AZN SS Equity), Abb Ltd-Reg (ABB SS Equity), Atlas Copco Ab-B Shs (ATCOB SS Equity), Atlas Copco Ab-A Shs (ATCOA SS Equity), Investor Ab-B Shs (INVEB SS Equity), Hennes & Mauritz Ab-B Shs (HMB SS Equity), Volvo Ab-B Shs (VOLVB SS Equity), Nordea Bank Abp (NDA SS Equity), Ericsson Lm-B Shs (ERICB SS Equity), Assa Abloy Ab-B (ASSAB SS Equity), Skandinaviska Enskilda Ban-A (SEBA SS Equity), Sandvik Ab (SAND SS Equity), Telia Co Ab (TELIA SS Equity), Svenska Handelsbanken-A Shs (SHBA SS Equity), Hexagon Ab-B Shs (HEXAB SS Equity), Swedbank Ab - A Shares (SWEDA SS Equity), Tele2 Ab-B Shs (TEL2B SS Equity), Alfa Laval Ab (ALFA SS Equity), Skanska Ab-B Shs (SKAB SS Equity), Skf Ab-B Shares (SKFB SS Equity), Electrolux Ab-Ser B (ELUXB SS Equity), Kinnevik Ab - B (KINVB SS Equity), Swedish Match Ab (SWMA SS Equity), Autoliv Inc-Swed Dep Receipt (ALIV SS Equity), Boliden Ab (BOL SS Equity), Svenska Cellulosa Ab Sca-B (SCAB SS Equity), Securitas Ab-B Shs (SECUB SS Equity), Getinge Ab-B Shs (GETIB SS Equity), Ssab Ab-A Shares (SSABA SS Equity).

Ciascuna delle due metaeuristiche è stata perciò utilizzata per ottimizzare due differenti portafogli, uno composto da titoli appartenenti al principale indice azionario italiano ed uno costituito da azioni appartenenti al principale indice svedese.

Prima di affrontare la fase computazionale è necessario illustrare il processo con cui sono stati implementati i due algoritmi. Le serie storiche dei prezzi scaricate da Bloomberg, come è già stato evidenziato, sono relative al periodo temporale che va dal 01/01/2015 al 30/06/2019. Questo periodo, però, è stato suddiviso in alcuni sottoperiodi *in-sample* (della durata di due anni), in cui avviene il calcolo dell'allocazione ottima di portafoglio e in altri *out-of-sample* (chiamati anche *futuro virtuale*) della durata di sei mesi, dove l'allocazione trovata nel periodo *in-sample* viene testata sui dati storici¹⁰ ai fini di verificare la capacità esplicativa del modello stesso. In questo modo si può verificare se il modello di portafoglio specificato consenta l'assunzione di decisioni d'investimento adeguate e la costruzione di portafogli performanti.

Per testare adeguatamente la bontà del modello si è ritenuto utile implementare quindi gli algoritmi in tre periodi temporali differenti, secondo la cosiddetta tecnica delle *rolling windows*¹¹:

- il primo periodo va dal 01/01/2015 al 31/12/2016 (*in-sample*) e dal 01/01/2017 al 30/06/2017 (*out-of-sample*);
- il secondo va dal 01/01/2016 al 31/12/2017 (*in-sample*) e dal 01/01/2018 al 30/06/2018 (*out-of-sample*);
- il terzo e ultimo periodo va dal 01/01/2017 al 31/12/2018 (*in-sample*) e dal 01/01/2019 al 30/06/2019 (*out-of-sample*).

¹⁰ Questa fase viene anche detta *backtesting*.

¹¹ Le *rolling windows*, in italiano *finestre mobili*, sono un tecnica che consente di analizzare una serie storica spostando di volta in volta l'analisi in un periodo temporale successivo, effettuando così la stima del modello su una porzione differente della serie storica stessa.

Alla luce di tutto ciò è possibile effettuare due tipologie di osservazioni. La prima è prettamente di tipo quantitativo, relativa alla differenza tra i due algoritmi sia in termini di tempo computazionale che in termini di qualità della soluzione ottima individuata, a parità di portafoglio ottimizzato. La seconda conclusione è invece di carattere economico/finanziario. Innanzitutto, si riesce a testare la bontà del modello sui dati storici e in scenari temporali differenti e poi, implementando lo stesso algoritmo nei due diversi mercati, si confrontano le composizioni azionarie dei portafogli ottimi costruiti su di essi. In questo modo si riesce a disporre anche di un utile strumento conoscitivo per trarre giudizi in merito all'economia nazionale dei Paesi stessi. I due indici azionari, infatti, essendo quelli più rappresentativi all'interno della stessa nazione, tendono a rispecchiare fedelmente l'andamento del tessuto societario nazionale e in generale dell'economia del Paese.

È necessario fare un'ulteriore precisazione. Ai fini dell'allocazione ottima di portafoglio, i due algoritmi devono innanzitutto calcolare i rendimenti logaritmici giornalieri a partire dalle serie storiche dei prezzi dei titoli sopra riportati. I rendimenti vengono perciò calcolati nella maniera seguente:

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

dove:

- $R_{i,t}$ è il rendimento logaritmico dell' i -esimo titolo al tempo t ;
- $P_{i,t}$ è il prezzo dell' i -esimo titolo al tempo t ;
- $P_{i,t-1}$ è il prezzo dell' i -esimo titolo al tempo $(t - 1)$.

5.3 Impostazione dei parametri del modello

Nel seguente paragrafo si illustrano i valori dei parametri che sono stati impostati relativamente ai vincoli, alla funzione obiettivo e alle metaeuristiche. Si suppone che tali valori rimangano costanti durante la fase computazionale e di seguito si riportano innanzitutto quelli relativi al sistema di vincoli e alla funzione obiettivo:

- numero minimo di titoli detenibili in portafoglio: $K_l = 5$;
- numero massimo di titoli detenibili in portafoglio: $K_u = 20$;
- percentuale d'investimento minima in ciascun titolo: $l_i = 2\%$;
- percentuale d'investimento massima in ciascun titolo: $u_i = 15\%$;
- il livello di confidenza $(1 - \alpha)$ dell'EVaR è supposto pari al 95% e quindi α è pari al 5%;
- il coefficiente di avversione al rischio z presente nella misura di rischio è supposto pari a 2.

Bisogna inoltre precisare che il rendimento obiettivo è stato posto pari alla media dei rendimenti attesi in-sample degli N titoli considerati per la stima campionaria:

$$\pi = \frac{1}{N} \sum_{i=1}^N \hat{r}_i.$$

Il rendimento medio giornaliero desiderato, quindi, dipende dalle serie storiche dei prezzi di ciascun titolo, usate per stimare il rendimento atteso del titolo i -esimo. Perciò, a seconda del periodo temporale considerato per la stima in-sample, si avranno differenti rendimenti desiderati. Il target return dipende allora dall'andamento del mercato in un determinato sottoperiodo campionario e quindi dall'economia di quel Paese.

5.3.1 PSO - impostazione dei parametri

Per risolvere il problema di ottimizzazione proposto si è deciso di utilizzare la versione della PSO con il peso d'inerzia. I parametri necessari per la determinazione della velocità V_i^{k+1} sono stati stabiliti in linea con quanto suggerito dalla letteratura e sono i seguenti:

- peso d'inerzia: $\omega = 0.7298$;
- coefficiente di accelerazione cognitivo: $\phi_1 = 1.49618$;
- coefficiente di accelerazione sociale: $\phi_2 = 1.49618$;
- numero di particelle¹²: 70;
- numero di iterazioni¹³: 1000.

A titolo esemplificativo si riporta in Figura 5.1 il grafico relativo alla convergenza della PSO all'ottimo, con un numero di iterazioni pari a 3000 e prendendo in considerazione i titoli dell'indice FTSE MIB. Si può notare che dopo 500 iterazioni il valore della fitness migliora marginalmente sino a stabilizzarsi.

¹² Normalmente per una buona esplorazione si richiede un numero di particelle almeno pari al doppio del numero dei titoli.

¹³ Si ritiene sufficiente un numero di iterazioni pari a 1000, in quanto un numero superiore di iterazioni comporta un tempo computazionale notevolmente maggiore, mentre si ha una riduzione solo marginale del valore della fitness.

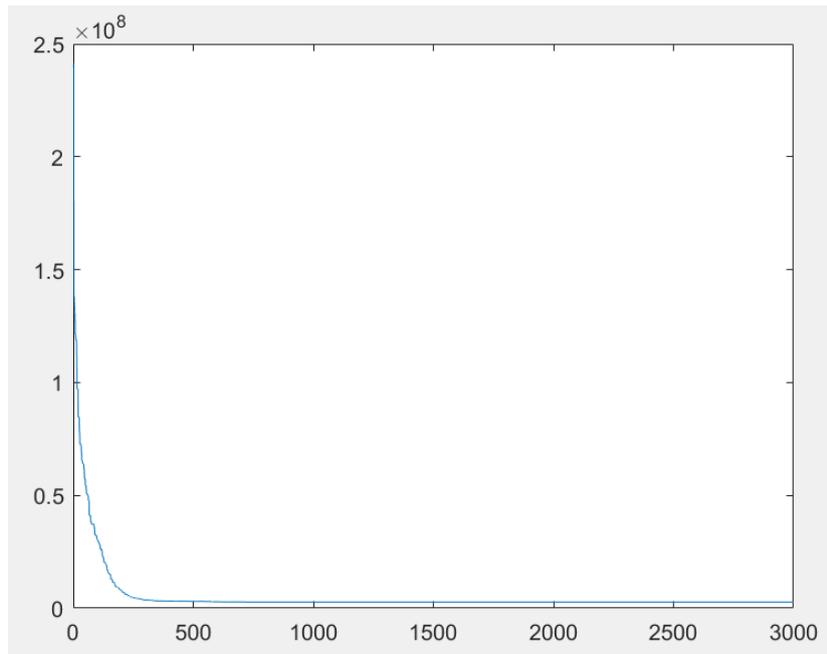


Figura 5.1 - Convergenza della PSO all'ottimo

L'ultimo parametro da stabilire è il valore di ε , necessario per la riformulazione del problema di selezione secondo il metodo delle penalità esatte. Questo parametro garantisce la corrispondenza tra la soluzione del problema originario vincolato e quella del problema non vincolato. Il corretto settaggio è quindi di fondamentale importanza. Sono stati effettuati dei test per diversi valori di ε , con un numero di iterazioni pari a 1000 e un numero di particelle pari a 70. Analizzando i valori medi normalizzati¹⁴ della funzione di fitness e la relativa deviazione standard per ogni valore di ε si è deciso di usare il valore di ε pari a 0.0000001, in quanto questo minimizza la fitness.

¹⁴ La normalizzazione della fitness si ottiene moltiplicando questa per il valore di ε .

5.3.2 FWA - impostazione dei parametri

Per quanto riguarda invece il FWA i parametri sono stati impostati nel seguente modo, come è suggerito dalla letteratura:

- ampiezza totale delle esplosioni: $\hat{A} = 40$;
- parametro che vincola il numero massimo di scintille: $b = 0.8$;
- parametro che vincola il numero minimo di scintille: $a = 0.04$;
- parametro che controlla il numero totale di scintille: $M_e = 40$;
- numero di scintille prodotte dall'esplosione Gaussiana per ciascun fuoco d'artificio: $M_g = 2$;
- numero di fuochi d'artificio¹⁵: 15;
- numero di iterazioni¹⁶: 1000.

Allo stesso modo, come è stato effettuato nella PSO, sono stati effettuati dei test per stabilire in valore del coefficiente di penalità e anche in questo caso è stato settato pari a 0.0000001.

5.4 Applicazione

Nel seguente paragrafo vengono infine illustrati i risultati ottenuti dall'implementazione delle due metaeuristiche. Come si è accennato in precedenza, si vogliono trarre due tipologie di conclusioni. Innanzitutto si effettua un confronto tra i due algoritmi, per verificare quale dei due consenta

¹⁵ Si è cercato di ottenere un numero totale di fuochi e scintille pari al numero di particelle nella PSO, così da rendere comparabili i due algoritmi.

¹⁶ In questo modo i due algoritmi sono comparabili e inoltre un numero maggiore di iterazioni fa aumentare eccessivamente il tempo computazionale, mentre il valore della fitness rimane stazionario.

di ottenere risultati migliori e in che tempo computazionale. Successivamente si effettuerà un confronto tra i portafogli costruiti sui due indici azionari, per verificare come vengono selezionati dalle due metaeuristiche.

5.4.1 Confronto tra PSO e FWA

Una delle prime osservazioni che salta subito all'occhio implementando le due metaeuristiche è quella relativa al grafico della convergenza della funzione di fitness all'ottimo globale del problema. Come si può notare dalla Figura 5.2, la convergenza avviene in maniera differente nei due algoritmi. Nella PSO, infatti, la capacità esplorativa delle particelle rende il processo di ricerca dell'ottimo più graduale e la funzione di fitness, che nel caso seguente ha raggiunto un valore sufficientemente basso già in prossimità dell'iterazione 400, diminuisce nelle successive iterazioni marginalmente ma in maniera "continua". Il FWA, invece, si comporta in maniera differente. La capacità esplorativa delle scintille prodotte dai fuochi d'artificio impedisce la convergenza "continua" alla soluzione ottima e vi è l'alternanza tra momenti di improvvisa diminuzione del valore delle fitness e di momenti di apparente stazionarietà.

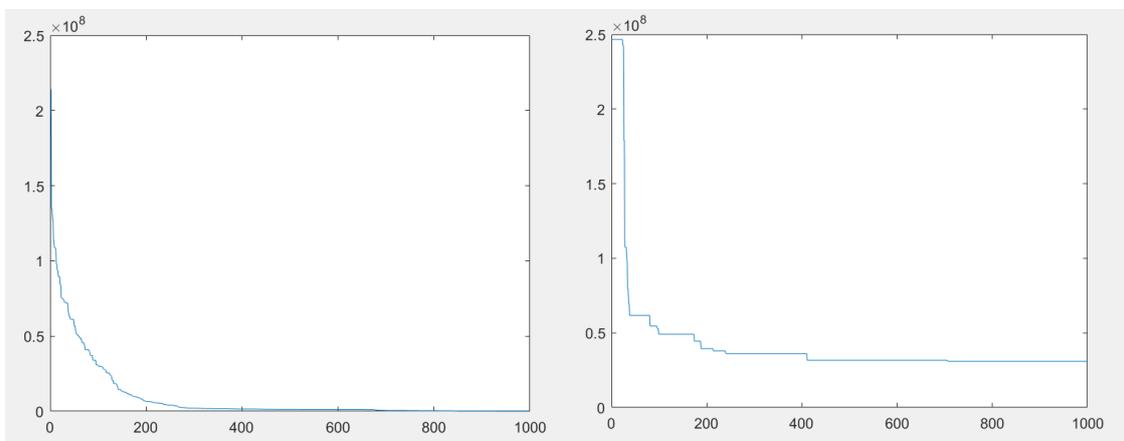


Figura 5.2 - Convergenza della funzione di fitness nella PSO (a sinistra) e nel FWA (a destra).

Un'altra notevole differenza tra i due algoritmi è il tempo computazionale necessario ad effettuare le K iterazioni. In tutti i periodi temporali ed i mercati, infatti, a parità di parametri, il FWA risulta sensibilmente più veloce della PSO. I risultati sono riportati nella Tabella 5.1.

	PSO	FWA
FTSE MIB periodo 1	302.38 secondi	81.64 secondi
FTSE MIB periodo 2	289.97 secondi	83.02 secondi
FTSE MIB periodo 3	292.77 secondi	82.73 secondi
OMX periodo 1	287.45 secondi	56.09 secondi
OMX periodo 2	287.02 secondi	56.30 secondi
OMX periodo 3	285.78 secondi	57.46 secondi

Tabella 5.1 - Tempo computazionale delle due metaeuristiche

5.4.2 Analisi degli scenari

In questo paragrafo vengono confrontati i due algoritmi in ogni scenario e per ciascun mercato, con attenzione alla violazione dei vincoli. Si ricorda che il primo scenario corrisponde al periodo che va dal 01/01/2015 al 31/12/2016 (in-sample) e dal 01/01/2017 al 30/06/2017 (out-of-sample). Il secondo va invece dal 01/01/2016 al 31/12/2017 (in-sample) e dal 01/01/2018 al 30/06/2018 (out-of-sample). Il terzo va infine dal 01/01/2017 al 31/12/2018 (in-sample) e dal 01/01/2019 al 30/06/2019 (out-of-sample).

5.4.2.1 Scenari FTSE MIB

Si riportano di seguito i risultati ottenuti con i due algoritmi nei tre differenti scenari per quanto riguarda l'indice azionario FTSE MIB.

	PSO	FWA
<i>Fitness</i>	65436.53	30984817.15
<i>Misura di rischio in-sample</i>	1.498	1.526
<i>Misura di rischio out-of-sample</i>	1.499	1.499
<i>Rendimento in-sample</i>	0.0000839	-0.0002096
<i>Rendimento out-of-sample</i>	0.00172	0.00155
<i>Violazione vincolo di bilancio</i>	0.00339	0.04527
<i>Violazione vincolo di rendimento atteso</i>	0.0000391	0.000331
<i>Violazione vincolo relativo al numero minimo di titoli</i>	0	0.97629
<i>Violazione vincolo relativo al numero massimo di titoli</i>	0	0
<i>Violazione vincolo di lower bound</i>	0.00003	0.334065
<i>Violazione vincolo di upper bound</i>	0.000279	0.009101
<i>Violazione vincolo di integralità</i>	0.006635	1.733406
<i>Numero titoli selezionati</i>	19.998	4.024
<i>Somma quote ottime</i>	1.00003	0.955

Tabella 5.3 - Output scenario 1 - FTSE MIB

Nella Tabella 5.3 sono riportati i risultati ottenuti dall'applicazione delle metaeuristiche al primo periodo temporale. Come è già stato sottolineato, nell'in-sample avviene la stima del portafoglio ottimo, che poi viene testato sui dati storici nell'out-of-sample. In questo scenario, i portafogli selezionati sia con la PSO che con il FWA consentirebbero all'ipotetico investitore che volesse effettuare l'investimento il 31/12/2016 di conseguire nella prima metà del 2017 rendimenti più elevati rispetto a quelli stimati. Questi rendimenti, prossimi allo zero nel periodo di stima e decisamente più elevati nell'out-of-sample, sono perfettamente spiegati dall'andamento dei mercati finanziari in quegli anni. Infatti nella prima metà del 2016 l'indice azionario italiano ha perso il 24.37% rispetto all'anno precedente, a causa del rallentamento della crescita cinese e dell'economia globale, del crollo dei prezzi del petrolio greggio, delle tensioni che hanno afflitto alcune banche italiane come MPS e degli effetti del referendum sulla "Brexit". Il 2017, invece, fu un anno completamente differente e l'indice FTSE MIB chiuse in rialzo di 2618.76 punti, pari al 13.59% di variazione positiva rispetto all'anno precedente. Nel corso di quest'anno, infatti, si osservò una riduzione del rapporto deficit/PIL, un aumento dell'inflazione che si trovava a livelli eccessivamente bassi e in generale un miglioramento dei conti pubblici.

Per quanto riguarda la rischiosità dei portafogli selezionati dalle due metaeuristiche, il FWA stima rispetto alla PSO un rischio maggiore nel periodo in-sample ma entrambi gli algoritmi costruiscono portafogli dotati della stessa rischiosità nel periodo del futuro virtuale. Quindi, poiché nell'out-of-sample la PSO consente di ottenere rendimenti più elevati, allora questa metaeuristica dovrebbe essere preferita da un investitore razionale in questo scenario temporale.

È interessante poi porre l'attenzione sulle violazioni dei vincoli. Si può infatti notare come la PSO rispetti sostanzialmente tutti i vincoli del modello, ad eccezione di alcuni casi in cui però commette violazioni trascurabili.

Nel caso del FWA, invece, si osserva un valore della fitness nettamente più elevato, a parità sia del numero di iterazioni che di tutti gli altri parametri. Questo avviene perché tale metaeuristica commette forti violazioni dei vincoli e in particolar modo del vincolo d'integralità, il che significa che l'algoritmo non è riuscito a selezionare un numero intero di titoli. Dai risultati si evince, infatti, che il FWA seleziona un numero di titoli pari a 4.02 e questo implica che, per ottenere il rendimento giornaliero di portafoglio pari a 0.16%, bisognerebbe poter acquistare alcuni titoli solo per una loro frazione, cosa che nella realtà non è possibile. La violazione del vincolo d'integralità può essere anche notata osservando che la somma delle quote ottime non è esattamente pari all'unità, cosa che avviene nel caso della PSO.

Esaminando poi il vincolo di rendimento atteso si può affermare che il rendimento giornaliero desiderato, calcolato come la media di tutti i rendimenti in-sample, viene sostanzialmente soddisfatto da entrambi gli algoritmi.

Un'ultima osservazione attiene al numero di titoli selezionati nel portafoglio. Dai risultati si evince che la PSO predilige una maggiore diversificazione, con un numero di titoli pari a 20, mentre il FWA è più restio a selezionare un elevato numero di titoli, prendendone in considerazione solo 4.

	PSO	FWA
<i>Fitness</i>	236351.09	44839510.43
<i>Misura di rischio in-sample</i>	1.498	1.513
<i>Misura di rischio out-of-sample</i>	1.498	1.498
<i>Rendimento in-sample</i>	0.00012	0.00033
<i>Rendimento out-of-sample</i>	0.0006	0.0008
<i>Violazione vincolo di bilancio</i>	0.0011	0.04
<i>Violazione vincolo</i>	0.0000053	0

<i>di rendimento atteso</i>		
<i>Violazione vincolo relativo al numero minimo di titoli</i>	0	0
<i>Violazione vincolo relativo al numero massimo di titoli</i>	0	0
<i>Violazione vincolo di lower bound</i>	0.02	0.496
<i>Violazione vincolo di upper bound</i>	0.00012	2.0043
<i>Violazione vincolo di integralità</i>	0.0034	1.943
<i>Numero titoli selezionati</i>	19.999	6.005
<i>Somma quote ottime</i>	1.000096	0.9598

Tabella 5.4 - Output scenario 2 - FTSE MIB

Nel secondo scenario, rappresentato in Tabella 5.4, si possono osservare in parte gli stessi risultati relativi al periodo temporale precedente. Nell'in-sample il FWA stima un portafoglio con un rischio ed un rendimento leggermente più elevati rispetto a quelli stimati dalla PSO.

Secondo entrambi gli algoritmi, inoltre, si può notare che i rendimenti conseguibili nel futuro virtuale sono più elevati rispetto a quelli stimati nell'in-sample. Questi maggiori rendimenti sono spiegati dall'andamento dei mercati finanziari nella prima metà del 2018, poiché, anche se complessivamente quest'anno è stato molto duro per i mercati finanziari, nel primo trimestre del 2018 il FTSE MIB ha avuto una serie di sedute di Borsa positive.

È necessario poi sottolineare che i portafogli selezionati dai due algoritmi hanno la stessa rischiosità nel futuro virtuale, ossia nella prima metà del 2018. In questo periodo, invece, il portafoglio a redditività maggiore è stato selezionato

dal FWA. Il 31/12/2017, perciò, un investitore razionale che analizza il profilo di rischio-rendimento degli investimenti, se volesse costruire un portafoglio sull'indice azionario FTSE MIB, propenderebbe per adottare la strategia d'investimento proposta dal FWA.

Per quanto riguarda le violazioni dei vincoli, infine, sono possibili osservazioni simili a quelle relative allo scenario precedente: la PSO riesce sempre a rispettare tutti i vincoli, mentre il FWA non rispetta né il vincolo di integralità, né quelli di lower e upper bound. Di conseguenza, oltre a non riuscire ad inserire in portafoglio un numero intero di titoli, il FWA seleziona i vari titoli violando le percentuali minime e massime di capitale totale, rispettivamente del 2% e del 15%, relative all'investimento in ciascun titolo del portafoglio.

Per quanto riguarda la numerosità dei titoli in portafoglio, infine, la PSO predilige sempre una maggiore diversificazione, mentre il FWA suggerisce di investire in un paniere di soli 3 titoli azionari italiani.

	PSO	FWA
<i>Fitness</i>	20582.42	35048951.76
<i>Misura di rischio in-sample</i>	1.498	1.513
<i>Misura di rischio out-of-sample</i>	1.499	1.499
<i>Rendimento in-sample</i>	0.00046	0.00027
<i>Rendimento out-of-sample</i>	0.0006	0.0013
<i>Violazione vincolo di bilancio</i>	0.00084	0.605
<i>Violazione vincolo di rendimento atteso</i>	0	0.000004
<i>Violazione vincolo relativo al numero minimo di titoli</i>	0	0
<i>Violazione vincolo</i>	0	0

<i>relativo al numero massimo di titoli</i>		
<i>Violazione vincolo di lower bound</i>	0.000004	0.175
<i>Violazione vincolo di upper bound</i>	0.00009	0.535
<i>Violazione vincolo di integralità</i>	0.0021	2.1896
<i>Numero titoli selezionati</i>	16.99	5.036
<i>Somma quote ottime</i>	1.00001	1.605

Tabella 5.5 - Output scenario 3 - FTSE MIB

Nel terzo e ultimo scenario, rappresentato in Tabella 5.5, si possono osservare risultati in linea con quanto detto per il periodo temporale precedente. Ciascun algoritmo consente un aumento del rendimento dal periodo di stima all'out-of-sample e i portafogli costruiti dai due algoritmi hanno la stessa rischiosità media nel futuro virtuale, periodo che coincide con la prima metà del 2019. Il FWA, però, consente di costruire portafogli a redditività superiore nell'out-of-sample e quindi questa metaeuristica risulta essere rispetto alla PSO una scelta migliore per un investitore razionale che voglia effettuare tale investimento il 31/12/2018. L'aumento della redditività dei portafogli nell'out-of-sample, oltre che indicatore di una buona selezione effettuata dai due algoritmi, è anche data dalla favorevole situazione economica del mercato italiano nei primi nove mesi del 2019. In questo periodo, infatti, discostandosi nettamente dal trend negativo di fine 2018, i mercati finanziari hanno registrato una ripresa generalizzata, provocata anche dalle iniezioni di liquidità e dai tagli dei tassi di interesse da parte delle Banche Centrali.

Per quanto riguarda le violazioni dei vincoli le osservazioni sono le medesime di quelle fatte con riferimento ai primi due scenari, in quanto si può notare una forte violazione del vincolo d'integralità da parte del FWA. Infine, si può notare che anche in questo caso la PSO predilige una diversificazione nettamente maggiore.

A titolo conclusivo si riportano in Appendice C le tabelle relative alle quote ottime d'investimento in ciascun titolo componente l'indice FTSE MIB, calcolate da entrambi gli algoritmi nei vari scenari.

Si comprende così quali sono secondo le due metaeuristiche i "migliori" titoli da inserire in portafoglio nei vari periodi temporali. A titolo esemplificativo si osservi che nel primo periodo la PSO suggerisce di investire maggiormente nella società di moda Salvatore Ferragamo; in particolare suggerisce un investimento pari al 14.54% del capitale disponibile. Questo significa che la PSO, dopo aver calcolato molte possibili combinazioni di titoli, è giunta alla conclusione che Salvatore Ferragamo, se inserito in portafoglio in quello specifico periodo, contribuisce a costruire dei portafogli con un miglior profilo di rischio-rendimento.

5.4.2.2 Scenari OMX Stockholm 30

Si riportano ora i risultati ottenuti nei tre diversi scenari per quanto riguarda la selezione di portafoglio sull'indice azionario OMX Stockholm 30.

	PSO	FWA
<i>Fitness</i>	412400.5	24865411.25
<i>Misura di rischio in-sample</i>	1.498	1.505
<i>Misura di rischio out-of-sample</i>	1.499	1.499
<i>Rendimento in-sample</i>	0.00024	0.000084

<i>Rendimento out-of-sample</i>	0.00099	0.00094
<i>Violazione vincolo di bilancio</i>	0.00048	0.309
<i>Violazione vincolo di rendimento atteso</i>	0	0.00005
<i>Violazione vincolo relativo al numero minimo di titoli</i>	0	1.266
<i>Violazione vincolo relativo al numero massimo di titoli</i>	0	0
<i>Violazione vincolo di lower bound</i>	0.0399	0.075
<i>Violazione vincolo di upper bound</i>	0.000008	0.0217
<i>Violazione vincolo di integralità</i>	0.0013	0.8121
<i>Numero titoli selezionati</i>	20.00001	3.733
<i>Somma quote ottime</i>	0.9999	0.6904

Tabella 5.6 - Output scenario 1 - OMX Stockholm 30

In Tabella 5.6 sono rappresentati i risultati relativi ai portafogli selezionati dai due algoritmi nel primo scenario. Si può notare come ciascuna metaeuristica selezioni portafogli che, nell'out-of-sample, siano dotati di redditività superiore rispetto a quella che è stata stimata nell'in-sample. Questo è in linea con il quadro macroeconomico svedese: la Svezia, infatti, ha registrato nel 2017 una crescita del 2.3%, inferiore all'anno precedente ma pur sempre sostenuta.

Per quanto riguarda la rischiosità, invece, si può osservare che la rischiosità dei portafogli selezionati dal FWA è inferiore nel periodo virtuale rispetto a quella

calcolata nella stima del modello, cosa che non avviene nel caso della PSO, dove le due rischiosità sono pressoché identiche.

In generale, entrambi gli algoritmi suggeriscono di investire in portafogli con profili di rischio-rendimento sostanzialmente molto simili. L'unica discriminante può essere vista nella redditività nell'out-of-sample del portafoglio generato dalla PSO, che risulta essere leggermente più elevata rispetto a quella del portafoglio selezionato dal FWA.

Per quanto riguarda la violazione dei vincoli si noti che, anche in questo caso, alcuni vincoli sono violati dal FWA, in particolar modo quello relativo al numero minimo di titoli e quello di integralità. Analogamente, come osservato precedentemente, il FWA non costruisce portafogli in linea con il principio della diversificazione.

	PSO	FWA
<i>Fitness</i>	4282.78	29180793.99
<i>Misura di rischio in-sample</i>	1.498	1.503
<i>Misura di rischio out-of-sample</i>	1.498	1.498
<i>Rendimento in-sample</i>	0.0001	0.0004
<i>Rendimento out-of-sample</i>	-0.00033	0.00036
<i>Violazione vincolo di bilancio</i>	0.000016	0.16
<i>Violazione vincolo di rendimento atteso</i>	0.00031	0.000045
<i>Violazione vincolo relativo al numero minimo di titoli</i>	0	0.194
<i>Violazione vincolo relativo al numero massimo di titoli</i>	0	0
<i>Violazione vincolo</i>	0.000001	0.2996

<i>di lower bound</i>		
<i>Violazione vincolo di upper bound</i>	0.0000022	0.0993
<i>Violazione vincolo di integralità</i>	0.00011	2.17
<i>Numero titoli selezionati</i>	12.00007	4.8058
<i>Somma quote ottime</i>	1.0000009	1.1593

Tabella 5.7 - Output scenario 2 - OMX Stockholm 30

In Tabella 5.7, che rappresenta il secondo scenario temporale, si nota che nel futuro virtuale entrambi i portafogli selezionati dai due algoritmi generano rendimenti inferiori rispetto a quelli che sono stati stimati nel modello, soprattutto nel caso della PSO, dove il rendimento medio out-of-sample risulta addirittura essere negativo. Questo rendimento negativo può essere spiegato dall'andamento macroeconomico globale. Dalla fine di gennaio 2018, infatti, i mercati finanziari hanno registrato una brusca decrescita causata da una serie di reazioni innescate dalla salita delle retribuzioni statunitensi; tale crescita ha infatti spaventato gli investitori, preoccupati che la Federal Reserve potesse essere indotta ad accelerare la politica di rialzo dei tassi con gravi conseguenze sui tassi del mercato obbligazionario e, di conseguenza, sull'intera crescita economica mondiale. Il dollaro ha così cominciato a rafforzarsi in modo significativo rispetto alle principali valute e tra le monete che si sono più deprezzate rispetto al dollaro figura appunto la corona svedese.

Tornando ai risultati forniti dagli algoritmi in questo scenario, è necessario evidenziare che il FWA seleziona portafogli con una rischiosità identica ma con un rendimento nettamente superiore a quelli selezionati con la PSO. Quindi, per un investitore che volesse effettuare un investimento in azioni svedesi il

31/12/2017, risulterebbe preferibile adottare la strategia d'investimento suggerita dal FWA.

Infine, anche in questo scenario si può notare che, mentre la PSO riesce ottimizzare il portafoglio nel rispetto dei vincoli imposti, il FWA commette delle forti violazioni, in particolar modo del vincolo di integralità.

	PSO	FWA
<i>Fitness</i>	273690.46	24574973.75
<i>Misura di rischio in-sample</i>	1.498	1.505
<i>Misura di rischio out-of-sample</i>	1.498	1.498
<i>Rendimento in-sample</i>	0.00015	0.000013
<i>Rendimento out-of-sample</i>	-0.00021	0.00091
<i>Violazione vincolo di bilancio</i>	0.0015	0.0479
<i>Violazione vincolo di rendimento atteso</i>	0	0.0001
<i>Violazione vincolo relativo al numero minimo di titoli</i>	0	0.9186
<i>Violazione vincolo relativo al numero massimo di titoli</i>	0	0
<i>Violazione vincolo di lower bound</i>	0.02	0.279
<i>Violazione vincolo di upper bound</i>	0.0003	0.169
<i>Violazione vincolo di integralità</i>	0.0067	1.043
<i>Numero titoli selezionati</i>	15.999	4.08
<i>Somma quote ottime</i>	1.0001	0.95

Tabella 5.8 - Output scenario 3 - OMX Stockholm 30

Nel terzo e ultimo scenario relativo all'indice azionario svedese, riportato in Tabella 5.8, figura una situazione analoga a quella precedente. I rendimenti prossimi allo zero o addirittura negativi, nel caso del portafoglio selezionato dalla PSO, possono trovare la loro spiegazione nell'andamento della corona svedese. La Svezia, nonostante negli ultimi quattro anni abbia riscontrato una crescita media annua del 3% e abbia un tasso di disoccupazione nettamente inferiore di quello italiano, sta in realtà attraverso una critica situazione economica, che trova il suo fondamento in un continuo indebolimento della corona svedese nei confronti di molte altre valute, come ad esempio l'euro e in un basso tasso di inflazione.

Tornando all'analisi dei risultati della Tabella 5.8 risulta infine che, come nel caso precedente, il Fireworks Algorithm sia da ritenere preferibile rispetto alla PSO dal punto di vista del profilo di rischio-rendimento, anche se però quest'algorithmo commette forti violazioni dei vincoli imposti nel modello.

Si vuole infine effettuare un confronto tra i portafogli costruiti sui due differenti indici azionari, a parità di ogni altra circostanza e dell'algorithmo utilizzato per l'ottimizzazione. Come è stato specificato all'inizio del capitolo, gli indici presi in considerazione sono costituiti dalle principali società operanti nei rispettivi Paesi e sono quindi degli indicatori del benessere economico del Paese stesso.

Osservando la rischiosità dei portafogli selezionati nei vari scenari dalle due metaeuristiche, risulta che il rischio misurato dall'Entropic Value-at-Risk relativamente ai portafogli azionari svedesi è, a parità di ogni altra circostanza, sempre minore del rischio misurato per i portafogli italiani. In altre parole è possibile costruire portafogli costituiti da azioni svedesi che siano meno rischiosi rispetto ai più sicuri portafogli composti da azioni italiane.

Questa minore rischiosità dei portafogli "svedesi" trova la sua spiegazione logica nel fatto che la Svezia ha una forte economia riconosciuta in tutto il mondo, tanto che la società di rating americana Standard & Poor's ha assegnato

a tale Paese il rating AAA, indicando così una situazione di perfetta solvibilità e di assenza del rischio creditizio. Il rating è, infatti, un indice di affidabilità creditizia e quindi di solidità e solvibilità del Paese.

L'Italia è invece in possesso del rating BBB, che indica una situazione economica più a rischio, che si ripercuote sul tessuto societario e di conseguenza anche sul mercato finanziario italiano.

A titolo conclusivo si riportano in Appendice D le tabelle relative alle quote ottime d'investimento in ciascun titolo componente l'indice azionario OMX Stockholm 30, calcolate da entrambi gli algoritmi nei tre differenti scenari temporali.

Conclusioni

Nel presente elaborato è stato costruito un modello di portafoglio e lo si è applicato a due differenti mercati azionari, uno relativo ad un Paese estremamente solido, come la Svezia ed uno relativo ad un Paese economicamente più debole, come l'Italia. Per la misurazione del rischio di portafoglio è stata usata una misura coerente ed entropica chiamata Entropic Value-at-Risk. Per la fase computazionale, invece, sono state implementate in Matlab le due metaeuristiche PSO e FWA. Si è potuto osservare che, a parte in qualche circostanza, il Fireworks Algorithm tende a selezionare portafogli che nel periodo out-of-sample generano rendimenti superiori, e hanno rischiosità pressoché identica, rispetto a quelli selezionati con la PSO. Secondo questi test, quindi, il FWA sembrerebbe essere una metaeuristica più valida. In realtà, come è stato osservato, quest'algoritmo non riesce a trovare soluzioni nei limiti dei vincoli che sono stati imposti nel modello, il che impedisce di trovare dei valori sufficientemente bassi della funzione di fitness. In particolare, il vincolo che viene maggiormente violato è quello di integralità e talvolta anche quello relativo al numero minimo di titoli da inserire in portafoglio. Inoltre, quest'algoritmo predilige una bassa diversificazione del portafoglio, a differenza della PSO che suggerisce portafogli sempre adeguatamente diversificati. Possiamo quindi affermare che il processo algoritmico del Fireworks Algorithm, che calcola inizialmente il numero delle scintille, l'ampiezza delle esplosioni e lo spostamento di ciascuna scintilla nello spazio delle soluzioni, non riesce ad esplorare e sfruttare efficientemente lo spazio di ricerca come avviene nella PSO. Nella PSO, infatti, grazie al movimento delle particelle determinato dall'aggiornamento continuo delle velocità, si riesce a trovare una soluzione ottima entro i limiti dello spazio ammissibile. Inoltre, l'ultima fase del FWA relativa alla selection strategy, dove vengono selezionate le scintille che diventeranno fuochi d'artificio nella successiva iterazione, forse

non garantisce la selezione delle scintille effettivamente “migliori”. Una possibile modifica al FWA potrebbe quindi consistere nell’aggiustare quest’ultima fase del processo algoritmico. Come è stato illustrato nel Paragrafo 4.6 relativamente alla strategia di selezione, l’algoritmo prevede la selezione della scintilla con la migliore posizione e di altre $(N - 1)$ scintille sulla base della distanza euclidea tra queste e le altre dello sciame. Si potrebbe aggiustare questa fase finale, ad esempio, imponendo all’algoritmo di selezionare le tre migliori scintille della popolazione corrente e di farle esplodere nella successiva iterazione; le altre $(N - 3)$ scintille potrebbero invece venire selezionate sulla base di un criterio che tenga in considerazione sia la distanza tra le altre, per garantire una buona esplorazione, che la propria funzione di fitness, per ridurre la probabilità di far esplodere scintille situate in una zona dello spazio poco promettente.

Appendice A

Codice Matlab: PSO.

```
clc;
clear all
close all
format long

%caricamento serie storiche e calcolo rendimenti logaritmici
[prezzi, testo] = xlsread('ftse');
[T,numvar] = size(prezzi);
rend = log(prezzi(2:end,:)./prezzi(1:end-1,:));

%suddivisione tra in-sample e out-of-sample
oos = 126; %numero rendimenti +1 non utilizzati per la stima (oos=out-of-sample)
ris = rend(1:T-oos,:); %rendimenti in-sample
ros = rend(T-oos+1:(T-1),:); %rendimenti out-of-sample

rm_is = mean(ris); %rendimenti medi in-sample
rm_oos = mean(ros); %rendimenti medi out-of-sample

%dati di input
Rp = mean(rm_is); %rendimento obiettivo giornaliero
Kd = 5; %numero minimo di titoli detenibili
Ku = 20; %numero massimo di titoli detenibili
d = ones(1,numvar)*0.02; %percentuale d'investimento minima in ciascun titolo
u = ones(1,numvar)*0.15; %percentuale d'investimento massima in ciascun titolo

%inizializzazione parametri PSO
P = 70; %numero particelle
c1 = 1.49618;
c2 = 1.49618;

a = 0.05; %confidenza
b = 2; %avversione al rischio
niter = 1000; %numero iterazioni
epsilon = 0.0000001;
iw = 0.7298; %peso d'inerzia

vmaxx = zeros(1,numvar);
vmaxz = zeros(1,numvar);

TT = (T-oos);

rho = zeros(P,1);
rm_rho = zeros(P,1);
R_is = zeros(P,1); %rendimento del portafoglio
```

```

R_oos = zeros(P,1);
vinc_1 = zeros(P,1); %vincolo di bilancio
vinc_2 = zeros(P,1); %vincolo sul rendimento medio minimo
vinc_3 = zeros(P,1); %vincolo sul numero minimo di titoli
vinc_4 = zeros(P,1); %vincolo sul numero massimo di titoli
app_5 = zeros(P,numvar);
vinc_5 = zeros(P,1); %vincolo sulla dimensione minima
app_6 = zeros(P,numvar);
vinc_6 = zeros(P,1); %vincolo sulla dimensione massima
app_7 = zeros(P,1);
vinc_7 = zeros(P,1); %z è 0 o 1
app_8 = zeros(P,numvar);
vinc_8 = zeros(P,1); %non negatività

%genera vettori posizioni/velocità e funzione obiettivo
x = rand(P,numvar);
vx = rand(P,numvar);
z = rand(P,numvar);
vz = rand(P,numvar);

f = ones(P,1)*1.0e+255;
x1 = zeros(P,numvar); %matrice di appoggio che tiene conto dell'effettiva presenza del titolo in
portafoglio

pbx = [x f]; %personal best
pbz = z;

gx = zeros(1,numvar+1); %gbest
gz = zeros(1,numvar);

differenza_is = zeros(TT,numvar);
prodotto_somma_is = zeros(TT,P);
differenza_oos = zeros(oos,numvar);
prodotto_somma_oos = zeros(oos,P);

tic

for k = 1:niter
    for i = 1:numvar
        vmaxx(i) = abs(max(x(:,i))-min(x(:,i)));
        vmaxz(i) = abs(max(z(:,i))-min(z(:,i)));
    end

    for p = 1:P

        for i = 1:numvar
            x1(p,i) = x(p,i)*z(p,i);
            app_5(p,i) = max(0,d(i)*z(p,i)-x(p,i));
            app_6(p,i) = max(0,x(p,i)-u(i)*z(p,i));
            app_7(p,i) = abs(z(p,i)*(1-z(p,i)));

```

```

    app_8(p,i) = max(0,-x1(p,i));
end

for t = 1:TT
    differenza_is(t,:) = ris(t:)-rm_is;
    prodotto_somma_is(t,p) = differenza_is(t,:)*x1(p,:);
end

for t = 1:(T-TT-1)
    differenza_oos(t,:) = ros(t:)-rm_oos;
    prodotto_somma_oos(t,p) = differenza_oos(t,:)*x1(p,:);
end

so = 0;
for conto = 1: TT
    so = so + exp(b*x1(p,1:end)*ris(conto,1:end));
end

rm_rho(p) = x1(p,:)*rm_is';
rho(p) = (1/b)*log(so/(a*TT));
vinc_1(p) = abs(sum(x1(p,:))-1); %vincolo di bilancio
vinc_2(p) = max(0,Rp-sum(x1(p,:)*rm_is')); %vincolo di redditività
vinc_3(p) = max(0,Kd-sum(z(p,:))); %vincolo numero minimo titoli
vinc_4(p) = max(0,(sum(z(p,:))-Ku)); %vincolo numero massimo titoli
vinc_5(p) = sum(app_5(p,:)); %vincolo percentuale minima
vinc_6(p) = sum(app_6(p,:)); %vincolo percentuale massima
vinc_7(p) = sum(app_7(p,:)); %vincolo d'integralità
vinc_8(p) = sum(app_8(p,:)); %vincolo di non negatività
end

%funzione obiettivo
f = rho+((1/epsilon)*(vinc_1+vinc_2+vinc_3+vinc_4+vinc_5+vinc_6+vinc_7));

%confronta il valore della funzione obiettivo con il pbest
for p = 1:P
    if f(p)<pbx(p,numvar+1)
        pbx(p,numvar+1) = f(p);
        for i = 1:numvar
            pbx(p,i) = x1(p,i);
            pbz(p,i) = z(p,i);
        end
    end
end
end

%identifica la particella con la posizione migliore
[minimo,posizione] = min(pbx(:,numvar+1));
gx(numvar+1) = minimo;

for i = 1:numvar

```

```

    gx(i) = pbx(posizione,i);
    gz(i) = pbz(posizione,i);
end

%cambia velocità e posizione
for p = 1:P
    for i = 1:numvar
        vx(p,i) = iw*vx(p,i)+c1*rand*(pbx(p,i)-x(p,i))+c2*rand*(gx(i)-x(p,i));
        vz(p,i) = iw*vz(p,i)+c1*rand*(pbz(p,i)-z(p,i))+c2*rand*(gz(i)-z(p,i));
        if vx(p,i)>vmaxx(i)
            vx(p,i) = vmaxx(i);
        end
        if vz(p,i)>vmaxz(i)
            vz(p,i) = vmaxz(i);
        end
        x(p,i) = x(p,i)+vx(p,i);
        z(p,i) = z(p,i)+vz(p,i);
    end
end
converg(k,:) = gx(:,end);
end

toc

quote_ottime = gx(1,1:numvar)'

so2 = 0;
for conto = 1:(T-TT-1)
    so2 = so2 + exp(b*ros*quote_ottime);
end

plot(converg)
best_fitness = gx(1,numvar+1)
rho_is = rho(posizione)
rho_oos = (1/b)*log(so2/(a*(T-TT-1)))
rm_out_of_sample = gx(1,1:numvar)*rm_oos'
rm_in_sample = gx(1,1:numvar)*rm_is'
n_titoli_selezionati = sum(z(posizione,:))
vinc_bilancio = vinc_1(posizione)
vinc_rendimento = vinc_2(posizione)
vinc_Kd = vinc_3(posizione)
vinc_Ku = vinc_4(posizione)
vinc_quota_min = vinc_5(posizione)
vinc_quota_max = vinc_6(posizione)
vinc_z_unitario = vinc_7(posizione)

```

Appendice B

Codice Matlab: FWA.

```
clear
clc
format long

% Caricamento serie storiche e calcolo rendimenti logaritmici
[prezzi,testo] = xlsread('ftse');
[T,numvar] = size(prezzi);
rend = log(prezzi(2:end,:)./prezzi(1:end-1,:));

%suddivisione tra in-sample e out-of-sample
oos = 126; %numero rendimenti +1 non utilizzati per la stima (oos = out-of-sample)
ris = rend(1:T-oos,:); %rendimenti in-sample
ros = rend(T-oos+1:(T-1),:); %rendimenti out-of-sample
rm_is = mean(ris); %rendimenti medi in-sample
rm_oos = mean(ros); %rendimenti medi out-of-sample
TT = T-oos;

%Dati di input
Rp = mean(rm_is); %rendimento obiettivo giornaliero
Kd = 5; %numero minimo di titoli detenibili
Ku = 20; %numero massimo di titoli detenibili
d = ones(1,numvar)*0.02; %percentuale d'investimento minima in ciascun titolo
u = ones(1,numvar)*0.15; %percentuale d'investimento massima in ciascun titolo
a = 0.05;
b = 2;

%parametri FWA
N = 15; %numero fuochi d'artificio
A_max = 40; %massima ampiezza del fuoco d'artificio
aa = 0.04; %costante per il numero minimo di scintille
bb = 0.8; %costante per il numero massimo di scintille
m = 40; %parametro che controlla il numero totale di scintille
epsilon = 0.0000001; %costante per evitare che il denominatore per il calcolo del numero di
scintille sia zero
m_gauss = 2; %numero di scintille prodotto dall'esplosione Gaussiana
niter = 1000; %numero d'iterazioni
x_max = 1000000000; %limite superiore dello spazio
x_min = -1000000000; %limite inferiore dello spazio
x_max_z=1000000000;
x_min_z=-1000000000;

%vincoli, vettori di appoggio per la funzione obiettivo
vinc_1 = zeros(N,1); %vincolo di bilancio
vinc_2 = zeros(N,1); %vincolo sul rendimento minimo atteso
```

```

app_1 = zeros(N,numvar); %matrice dei pesi di ciascun titolo i per ciascuna particella p
vinc_3 = zeros(N,1); %vincolo di non negatività
vinc_4 = zeros(N,1); %vincolo sul numero minimo di titoli
vinc_5 = zeros(N,1); %vincolo sul numero massimo di titoli
app_2 = zeros(N,numvar);
vinc_6 = zeros(N,1); %z è 0 o 1
app_3 = zeros(N,numvar);
app_4 = zeros(N,numvar);
vinc_7 = zeros(N,1); %vincolo frazione minima
vinc_8 = zeros(N,1); %vincolo frazione massima
differenza_is = zeros(TT,numvar);
prodotto_somma_is = zeros(TT,N);
differenza_oos = zeros(TT-TT-1,numvar);
prodotto_somma_oos = zeros(T-TT-1,N);
rho = zeros(N,1);
x1 = zeros(N,numvar);

s = zeros(N,1);
x = rand(N,numvar); %inizializzazione delle posizioni delle esplosioni
z = rand(N,numvar);
gx = zeros(1,numvar); %miglior funzione di fitness
gz = zeros(1,numvar);
A = zeros(N,1);
rm_rho = zeros(N,1);

for n=1:N
    for i=1:numvar
        x1(n,i)=x(n,i)*z(n,i);
        app_1(n,i)=max(0,-x1(n,i));
        app_2(n,i)=abs(z(n,i)*(1-z(n,i)));
        app_3(n,i)=max(0,d(i)*z(n,i)-x(n,i));
        app_4(n,i)=max(0,x(n,i)-u(i)*z(n,i));
    end
    for t=1:TT
        differenza_is(t,:)=ris(t,:)-rm_is;
        prodotto_somma_is(n,t)=differenza_is(t,:)*x1(n,:);
    end
    for t=1:(T-TT-1)
        differenza_oos(t,:)=ros(t,:)-rm_oos;
        prodotto_somma_oos(n,t)=differenza_oos(t,:)*x1(n,:);
    end

    so=0;
    for conto = 1:TT
        so = so + exp(b*x1(n,1:end)*ris(conto,1:end));
    end

    rm_rho(n)=x1(n,:)*rm_is';
    rho(n)= (1/b)*log(so/(a*TT));
    vinc_1(n)=abs(sum(x1(n,:))-1); %vincolo di bilancio
    vinc_2(n)=max(0,Rp-sum(x1(n,:)*rm_is')); %vincolo di redditività
end

```

```

vinc_3(n)=sum(app_1(n,:)); %vincolo non negativita
vinc_4(n)=max(0,Kd-sum(z(n,:))); %vincolo di cardianlità
vinc_5(n)=max(0,sum(z(n,:))-Ku); %vincolo di cardinalità
vinc_6(n)=sum(app_2(n,:)); %vincolo d'integralità
vinc_7(n)=sum(app_3(n,:)); %vincolo di frazione minima
vinc_8(n)=sum(app_4(n,:)); %vincolo di frazione massima
end
misura_di_rischio=rho;
vinc_1_a=vinc_1;
vinc_2_a=vinc_2;
vinc_3_a=vinc_3;
vinc_4_a=vinc_4;
vinc_5_a=vinc_5;
vinc_6_a=vinc_6;
vinc_7_a=vinc_7;
vinc_8_a=vinc_8;

%funzione obiettivo
f=misura_di_rischio+((1/epsilon)*(vinc_1+vinc_2+vinc_4+vinc_5+vinc_6+vinc_7+vinc_8));
tic
for k=1:niter
    for n=1:N
        s(n)=round( m*(((max(f)-f(n))+epsilon)/((sum(max(f)-f)+epsilon))));
        if s(n)<aa*m
            s(n)=round(aa*m);
        end
        if s(n)>bb*m
            s(n)=round(bb*m);
        end
        A(n)=A_max*((f(n)-min(f)+epsilon)/((sum(f-min(f))+epsilon)));
        if A(n)<0.5
            A(n)=0.5;
        end
    end
    sparks=s;
    ampiezza=A;
    tot=sum(sparks);
    new=zeros(tot,numvar);
    new_z=zeros(tot,numvar);
    matrice_A=A.*ones(N,numvar); %matrice che contiene il vettore delle ampiezze concatenato
    per numvar volte
    h=zeros(tot,numvar); %vettore di appoggio per il displacement operation
    h_z=zeros(tot,numvar);
    it=1; %permette di raccogliere gli elementi in una matrice nel ciclo for successivo
    for i=1:numvar
        for n=1:N
            LL=ones(s(n),numvar); %matrice di appoggio che cambia la dimensione delle righe ad
            ogni iterazione, di dimensioni s(n) x numvar, con s(n) numero di scintille prodotto da ogni
            fireworks n
            CC=LL*matrice_A(n,i); %matrice che somma la matrice di zeri LL alla matrice_A in ogni
            riga

```

[temp,of]=size(CC) ; %dimensioni della matrice CC ad ogni iterazione: le dimensioni della matrice CC cambiano ad ogni iterazione in quanto il numero di scintille perodotte da ogni fuoco d'artificio è diverso

```

for t=1:temp
    h(it,i)=A(n)*rand ; %calcola lo spostamento
    h_z(it,i)=A(n)*rand;%(A(n)*rand );
    new(it,i)=x(n,i)+h(it,i); %calcola le posizioni delle scintille
    new_z(it,i)=z(n,i)+h_z(it,i);
    it=it+1; %permette di raccogliere gli elementi in una matrice
end
end
end
spostamento=h;
spostamento_z=h_z;
mat=new ;
%la matrice mat è una matrice di dimensioni (tot*numvar) x numvar che include ad ogni
colonna un numero di zeri pari a tot*(numvar-1),
%per togliere gli zeri si segue il seguente procedimento
mat_z=new_z;
[jj]=size(mat);
[jj_z]=size(mat_z);
mat=mat';
mat_z=mat_z';
mat(mat==0) = []; %comando che toglie gli zeri
mat_z(mat_z==0) = [];
mat_dim_corrette=reshape(mat, jj(1)-tot*(numvar-1), jj(2)); %matrice con le posizioni delle
scintille di dimensioni tot x numvar
mat_dim_corrette_z=reshape(mat_z, jj_z(1)-tot*(numvar-1), jj_z(2));
d=randi([0,1],1,numvar);
mat_z_dim=mat_dim_corrette.*d;
mat_z_dim_z=mat_dim_corrette_z.*d;
% il seguente passaggio serve per aggiungere alla matrice di z dimensioni le dimemnsioni che
sono state tolte per riportare la matrice alla dimensioni tot x numvar; le dimensioni che non
sono state modificate rimangono infatti uguali a quelle della matrice x, solo che devono essere
replicate per tutte le scintille appartenenti ad ogni fireworks della matrice x
new_1=zeros(tot,numvar);
new_1_z=zeros(tot,numvar);
et=1;
for i=1:numvar
    for n=1:N
        MM=zeros(s(n),numvar); %matrice di appoggio che cambia la dimensione delle righe ad
ogni iterazione, di dimensioni s(n) x numvar, con s(n) numero di scintille prodotto da ogni
fireworks n
        DD=MM+matrice_A(n,i); %matrice che somma la matrice di zeri LL alla matrice_A in
ogni riga
        [temp_1,of_1]=size(DD); %dimensioni della matrice CC ad ogni iterazione: le dimensioni
della matrice CC cambiano ad ogni iterazione in quanto il numero di scintille perodotte da ogni
fuoco d'artificio è diverso
        for t=1:temp_1
            new_1(et,i)=x(n,i); %calcola le posizioni delle scintille
            new_1_z(et,i)=z(n,i);

```

```

        et=et+1; %permette di raccogliere gli elementi in una matrice
    end
end
end
mat_1=new_1;
mat_1_z=new_1_z;
jj_1=size(mat_1);
jj_1_z=size(mat_1_z);
mat_1=mat_1';
mat_1_z=mat_1_z';
mat_1(mat_1==0) = []; %comando che toglie gli zeri
mat_1_z(mat_1_z==0) = [];
mat_dim_corrette_1=reshape(mat_1,jj_1(1)-tot*(numvar-1),jj_1(2));
mat_dim_corrette_1_z=reshape(mat_1_z,jj_1_z(1)-tot*(numvar-1),jj_1_z(2));

%mapping rule
for i=1:numvar
    if mat_z_dim(:,i)<0.00000000000000000001
        mat_z_dim(:,i)=mat_dim_corrette_1(:,i);
        if mat_z_dim_z(:,i)<0.00000000000000000001
            mat_z_dim_z(:,i)=mat_dim_corrette_1_z(:,i);
        end
    end
end
end
total_sparks_after_displacement=mat_z_dim;
total_sparks_after_displacement_z=mat_z_dim_z;
d_g=randi([0,1],1,numvar);
g=abs(normrnd(1,1,[m_gauss*N,1]));
g_z=abs(normrnd(1,1,[m_gauss*N,1]));
g_1=g.*ones(m_gauss*N,numvar);
g_1_z=g_z.*ones(m_gauss*N,numvar);
LLP=repmat(x,m_gauss,1);
LLP_z=repmat(z,m_gauss,1);
new_g=LLP.*g_1;
new_g_z=LLP_z.*g_1_z;
sparks_after_gaussian=new_g.*d_g;
sparks_after_gaussian_z=new_g_z.*d_g;
for i=1:numvar
    if sparks_after_gaussian(:,i)<0.00000000000000000001
        sparks_after_gaussian(:,i)=LLP(:,i);
        if sparks_after_gaussian_z(:,i)<0.00000000000000000001
            sparks_after_gaussian_z(:,i)=LLP_z(:,i);
        end
    end
end
end
total_sparks_after_gaussian=sparks_after_gaussian;
total_sparks_after_gaussian_z=sparks_after_gaussian_z;
%sparks_after_gaussian=new_g(:,1:z_dimensions);
%sparks_after_gaussian_z=new_g_z(:,1:z_dimensions);
%total_sparks_after_gaussian=[sparks_after_gaussian,LLP(:,z_dimensions+1:end)];
%total_sparks_after_gaussian_z=[sparks_after_gaussian_z,LLP_z(:,z_dimensions+1:end)];

```

```

%MATRICE CON TUTTE LE SCINTILLE SIA DELL'ESPLOSIONE DEL PRIMO TIPO
%(DISPLACEMENT OPERATION) E SIA DELL'ESPLOSIONE DI SECONDO TIPO
(GAUSSIAN
%OPERATION
matrice_all_sparks=cat(1,total_sparks_after_displacement,total_sparks_after_gaussian);
matrice_all_sparks_z=cat(1,total_sparks_after_displacement_z,
total_sparks_after_gaussian_z);
%mapping rule
[tot_sparks, numvar]=size(matrice_all_sparks);
for ss=1:tot_sparks
    for i=1:numvar
        if matrice_all_sparks(ss,i)>x_max
            matrice_all_sparks(ss,i)=abs(matrice_all_sparks(ss,i))-x_max;
        end
        if matrice_all_sparks(ss,i)<x_min
            matrice_all_sparks(ss,i)=abs(matrice_all_sparks(ss,i))-x_max;
        end
        if matrice_all_sparks_z(ss,i)>x_max_z
            matrice_all_sparks_z(ss,i)=abs(matrice_all_sparks_z(ss,i))-x_max_z;
        end
        if matrice_all_sparks_z(ss,i)<x_min_z
            matrice_all_sparks_z(ss,i)=abs(matrice_all_sparks_z(ss,i))-x_max_z;
        end
    end
end
vinc_1_b=zeros(tot_sparks,1) ; %vincolo di bilancio
vinc_2_b=zeros(tot_sparks,1); %vincolo sul rendimento minimo atteso
app_1_b=zeros(tot_sparks,numvar); %matrice dei pesi di ciascun titolo i per ciascuna
particella p
vinc_3_b=zeros(tot_sparks,1); %vincolo di non negatività
vinc_4_b=zeros(tot_sparks,1); %vincolo sul numero minimo di titoli
vinc_5_b=zeros(tot_sparks,1); %vincolo sul numero massimo di titoli
app_2_b=zeros(tot_sparks,numvar);
vinc_6_b=zeros(tot_sparks,1); %z è 0 o 1
app_3_b=zeros(tot_sparks,numvar);
app_4_b=zeros(tot_sparks,numvar);
vinc_7_b=zeros(tot_sparks,1); %vincolo frazione minima
vinc_8_b=zeros(tot_sparks,1); %vincolo frazione massima
R_ss=zeros(tot_sparks,1);
rho_b=zeros(tot_sparks,1);
x1_b=zeros(tot_sparks,numvar);
%%scelta f(x) minore (gbest)
for ss=1:tot_sparks
    for i=1:numvar
        x1_b(ss,i)=matrice_all_sparks(n,i)*matrice_all_sparks_z(n,i);
        app_1_b(ss,i)=max(0,-x1_b(ss,i));
        app_2_b(ss,i)=abs(matrice_all_sparks_z(ss,i)*(1-matrice_all_sparks_z(ss,i)));
        app_3_b(ss,i)=max(0,d(i)*matrice_all_sparks_z(ss,i)-matrice_all_sparks(ss,i));
        app_4_b(ss,i)=max(0,matrice_all_sparks(ss,i)-u(i)*matrice_all_sparks_z(ss,i));
    end
    R_ss(ss)=x1_b(ss,:)*rm_is';
end

```

```

    rho_b(ss)=(a/tot_sparks)*(sum(max(0,matrice_all_sparks(ss,:)))+(1-
a)*(((1/tot_sparks)*(sum((max(0,-matrice_all_sparks(ss,:)).^b)))^(1/b))-R_ss(ss));
    vinc_1_b(ss)=abs(sum(x1_b(ss,:))-1);
    vinc_2_b(ss)=max(0,Rp-sum(x1_b(ss,:)*rm_is'));
    vinc_3_b(ss)=sum(app_1_b(ss,:));
    vinc_4_b(ss)=max(0,Kd-sum(matrice_all_sparks_z(ss,:)));
    vinc_5_b(ss)=max(0,sum(matrice_all_sparks_z(ss,:))-Ku);
    vinc_6_b(ss)=sum(app_2_b(ss,:));
    vinc_7_b(ss)=sum(app_3_b(ss,:));
    vinc_8_b(ss)=sum(app_4_b(ss,:));
end

f_b=rho_b+((1/epsilon)*(vinc_1_b+vinc_2_b+vinc_4_b+vinc_5_b+vinc_6_b+vinc_7_b+vinc_8_b));
all_position=cat(1, matrice_all_sparks, x);
dist_app=zeros(tot_sparks+N,tot_sparks+N);
R_app=zeros(1, tot_sparks);
for j=1:tot_sparks+N
    for p=1:tot_sparks+N
        dist_app(p,j)=(sqrt(sum((all_position(j,:)-all_position(p,:)).^2)));
    end
end
distanza=dist_app;
R=sum(distanza);
tot_R=sum(R);
prob=R/tot_R;
R_1=R';
%selection_best_sparks
all_z=cat(1,matrice_all_sparks_z,z);
all_fitness=cat(1,f_b,f);
all_rho=cat(1,rho_b,misura_di_rischio);
all_vinc_1=cat(1,vinc_1_b,vinc_1_a);
all_vinc_2=cat(1,vinc_2_b,vinc_2_a);
all_vinc_3=cat(1,vinc_3_b,vinc_3_a);
all_vinc_4=cat(1,vinc_4_b,vinc_4_a);
all_vinc_5=cat(1,vinc_5_b,vinc_5_a);
all_vinc_6=cat(1,vinc_6_b,vinc_6_a);
all_vinc_7=cat(1,vinc_7_b,vinc_7_a);
all_vinc_8=cat(1,vinc_8_b,vinc_8_a);
all_x1=cat(1,x1_b,x1);
matrice_tot=[all_position, all_fitness,all_z, all_x1, R_1, all_rho, all_vinc_1, all_vinc_2,
all_vinc_3, all_vinc_4, all_vinc_5, all_vinc_6, all_vinc_7, all_vinc_8];
ordine_minima_fitness=sortrows(matrice_tot,numvar+1);
selection_best_spark=ordine_minima_fitness(1,:);
%selection_sparks
matrice_senza_best_spark=ordine_minima_fitness(2:end,:);
ordine_massima_distanza=sortrows(matrice_senza_best_spark,
numvar+1+numvar+numvar+1);
selection_sparks=ordine_massima_distanza(end-N+2:end,:);%(1:N-1,:)
selection_sparks_for_next_iteration=cat(1, selection_best_spark,selection_sparks);
selection_x=selection_sparks_for_next_iteration(:,1:numvar);
selection_fitness=selection_sparks_for_next_iteration(:,numvar+1);

```

```

selection_z=selection_sparks_for_next_iteration(:,numvar+1+1:numvar+numvar+1);
selection_x1=selection_sparks_for_next_iteration(:, numvar+1+numvar+1:end-10);
selection_R_1=selection_sparks_for_next_iteration(:,end-9);
selection_rho=selection_sparks_for_next_iteration(:,end-8);
selection_vinc_1_a=selection_sparks_for_next_iteration(:,end-7);
selection_vinc_2_a=selection_sparks_for_next_iteration(:,end-6);
selection_vinc_3_a=selection_sparks_for_next_iteration(:,end-5);
selection_vinc_4_a=selection_sparks_for_next_iteration(:,end-4);
selection_vinc_5_a=selection_sparks_for_next_iteration(:,end-3);
selection_vinc_6_a=selection_sparks_for_next_iteration(:,end-2);
selection_vinc_7_a=selection_sparks_for_next_iteration(:,end-1);
selection_vinc_8_a=selection_sparks_for_next_iteration(:,end);
f=selection_fitness;
x=selection_x;
z=selection_z;
x1=selection_x1;
misura_di_rischio=selection_rho;
vinc_1_a=selection_vinc_1_a;
vinc_2_a=selection_vinc_2_a;
vinc_3_a=selection_vinc_3_a;
vinc_4_a=selection_vinc_4_a;
vinc_5_a=selection_vinc_5_a;
vinc_6_a=selection_vinc_6_a;
vinc_7_a=selection_vinc_7_a;
vinc_8_a=selection_vinc_8_a;
gx= f(1);
converg(k,:)=gx(:,end);
end
toc
plot(converg)

```

%RISULTATI

```

f(1)
quote_otime=x1(1,:)'
so2 = 0;
for conto = 1:(T-TT-1)
    so2 = so2 + exp(b*ros*quote_otime);
end
rho_oos = (1/b)*log(so2/(a*(T-TT-1)))
rho_is=rho(1,1)
rm_out_of_sample = x1(1,1:numvar)*rm_oos'
rm_in_sample = x1(1,1:numvar)*rm_is'
n_titoli_selezionati=sum(z(1,:))
vinc_bilancio=vinc_1_a(1)
vinc_redditivita=vinc_2_a(1)
vinc_numero_min=vinc_4_a(1)
vinc_numero_max=vinc_5_a(1)
vinc_integralita=vinc_6_a(1)
vinc_frazione_minima=vinc_7_a(1)
vinc_frazione_massima=vinc_8_a(1)

```

Appendice C

Quote ottime d'investimento in ciascun titolo componente il FTSE MIB nei vari scenari.

FTSE - scenario 1		
	PSO	FWA
A2A	0.026641256351480	0.000000005404993
Amplifon	0.024639558539823	0.000000008232750
Atlantia	-0.000000000000000	0.000016488603559
Azimut Holding	0.028476685885025	0.065103825644206
Banco BPM	0.089103019272987	0.072607833545671
BPER Banca	0.000000000000000	0.007993432928794
Buzzi Unicem	0.049452099462751	0.014524567518914
Campari	0.057942337254487	0.000043149054259
CNH Industrial	0.069472331193195	0.174810723561179
DiaSorin	0.028111409761832	0.002631484860072
Enel	0.067487066470825	0.000000006786184
Eni	0.000000000000000	0.000002439619495
Exor	0.020259651515949	0.000016545366722
Fiat Chrysler Automobiles	0.028418142630389	0.017236413077535
FinecoBank	0.026317038767547	0.000000003896515
Generali	0.000000000000000	0.030556816402865
Hera	0.000000000000000	0.127538607863022
Intesa Sanpaolo	0.000000000000000	0.185795243013967
Juventus Football Club	0.000000000000000	0.016582811449257
Leonardo	0.043731064998146	0.000001030624450
Mediobanca	0.028789728039834	0.013085006508016
Moncler	0.083508964062914	0.010797308620242
Prysmian	-0.000000000000000	0.000000096031941
Recordati	0.000000000000000	0.000000212431514
Saipem	-0.000000000000000	0.000000199330674
Salvatore Ferragamo	0.145360683686780	0.043824269929206
Snam	0.070304761554260	0.000211303075691
STMicroelectronics	0.023286614614631	0.020416283886343
Telecom Italia	0.025621788482394	0.078405846839266
Tenaris	-0.000000000000000	0.001652366112136
Terna - Rete Elettrica Nazionale	0.000000000000000	0.000020217655457
UBI Banca	0.063109493210569	0.033548420649965
UniCredit	-0.000000000000000	0.036369708199158
Unipol	-0.000000000000000	0.000929007666472
UnipolSai Assicurazioni	0.000000000000000	0.000000245362937

FTSE - scenario 2		
	PSO	FWA
A2A	-0.0000000000000000	0.122950533973416
Amplifon	0.137416473374456	0.115560332717622
Atlantia	-0.0000000000000000	0.000000922135296
Azimut Holding	0.021308856952352	0.000000211486608
Banco BPM	0.022525557832192	0.000000039261982
BPER Banca	0.0000000000000000	0.000000003933693
Buzzi Unicem	0.0000000000000000	0.000000065824976
Campari	0.051320329401502	0.000000000527714
CNH Industrial	0.000000038700599	0.000000026806211
DiaSorin	-0.0000000000000000	0.022671344833476
Enel	0.033712726379034	0.000000004145781
Eni	-0.0000000000000000	0.000079848137482
Exor	-0.0000000000000000	0.176311643253964
Fiat Chrysler Automobiles	-0.0000000000000000	0.000000002293347
FinecoBank	0.026358082204707	0.000000000043470
Generali	0.031263136754925	0.083768728810148
Hera	0.0000000000000000	0.000000000051807
Intesa Sanpaolo	0.031339775532271	0.001127848559041
Juventus Football Club	0.022337840472904	0.000000000015514
Leonardo	0.109384023115359	0.000000000693661
Mediobanca	0.024272027108832	0.000000007994399
Moncler	0.145316025675051	0.082479408521522
Prysmian	-0.0000000000000000	0.000000013902134
Recordati	0.0000000000000000	0.050780184074749
Saipem	0.022339103733701	0.000000012319039
Salvatore Ferragamo	0.025501682020788	0.091754534499495
Snam	-0.0000000000000000	0.007087075799401
STMicroelectronics	0.0000000000000000	0.000000000480611
Telecom Italia	-0.0000000000000000	0.000000038824960
Tenaris	-0.0000000000000000	0.000000000000261
Terna - Rete Elettrica Nazionale	0.064140213208762	0.000000000971713
UBI Banca	0.034615197952537	0.000003195482318
UniCredit	0.134844204775269	0.000000005481707
Unipol	0.038758143009754	0.000000008703980
UnipolSai Assicurazioni	0.023342850857601	0.205181156961282

FTSE - scenario 3		
	PSO	FWA
A2A	0.0000000000000000	0.004798411484051
Amplifon	-0.0000000000000000	0.000603122787646
Atlantia	0.095592915501811	0.000030632296907
Azimut Holding	-0.0000000000000000	0.000218076353376
Banco BPM	-0.0000000000000000	0.000019042949248
BPER Banca	-0.0000000000000000	0.015282131719339
Buzzi Unicem	-0.0000000000000000	0.146167272838589
Campari	0.150146585392621	0.000002538428849
CNH Industrial	-0.0000000000000000	0.000000089260731
DiaSorin	0.020685930427726	0.010148967382644
Enel	0.0000000000000000	0.000000007784091
Eni	-0.0000000000000000	0.007621093797098
Exor	0.135294336503604	0.000030644838112
Fiat Chrysler Automobiles	0.023417858323728	0.000195461076908
FinecoBank	0.0000000000000000	0.000009626376508
Generali	0.0000000000000000	0.195887530426861
Hera	-0.0000000000000000	0.057252771051653
Intesa Sanpaolo	-0.0000000000000000	0.000016336645793
Juventus Football Club	0.024630651723924	0.024038385349855
Leonardo	0.021781938433957	0.178994294992227
Mediobanca	-0.0000000000000000	0.267352022311878
Moncler	0.123978771246138	0.169911214054832
Prysmian	0.035615982995109	0.078893602618607
Recordati	0.054451292314857	0.000004253802493
Saipem	0.094092868370608	0.000000137772680
Salvatore Ferragamo	0.022257496843119	0.074093064206269
Snam	0.022581424383611	0.181621910488644
STMicroelectronics	-0.0000000000000000	0.000001213606587
Telecom Italia	0.050824806103324	0.000000574142664
Tenaris	0.0000000000000000	0.000000009512551
Terna - Rete Elettrica Nazionale	0.024308539194996	0.125881768337699
UBI Banca	0.071573921369425	0.015420193294644
UniCredit	0.0000000000000000	0.002394350471447
Unipol	0.028776752303838	0.048111517823458
UnipolSai Assicurazioni	-0.0000000000000000	0.000062490776131

Appendice D

Quote ottime d'investimento in ciascun titolo componente l'OMX Stockholm 30 nei vari scenari.

OMX - scenario 1		
	PSO	FWA
Astrazeneca Plc	0.038693053797817	0.044816958355483
Abb Ltd-Reg	0.000000000000000	0.00000000238859
Atlas Copco Ab-B Shs	0.050098556476136	0.113632783080648
Atlas Copco Ab-A Shs	0.023974599218899	0.000005693107880
Investor Ab-B Shs	0.034832693338956	0.000002470623469
Hennes & Mauritz Ab-B Shs	0.025827878529644	0.000016349851114
Volvo Ab-B Shs	0.067558511719518	0.057430980345345
Nordea Bank Abp	-0.000000000000000	0.000533338517688
Ericsson Lm-B Shs	0.024888157574587	0.000078901572556
Assa Abloy Ab-B	-0.000000000000000	0.00000000957059
Skandinaviska Enskilda Ban-A	0.053165194081717	0.000274589403414
Sandvik Ab	0.035903187740154	0.000115094190407
Telia Co Ab	0.000000000000000	0.065648051260991
Svenska Handelsbanken-A Shs	-0.000000000000000	0.000090311176942
Hexagon Ab-B Shs	0.069393117298314	0.061736823357256
Swedbank Ab-A Shares	-0.000000332438835	0.000478801232945
Tele2 Ab-B Shs	0.000000000000000	0.00000000514509
Alfa Laval Ab	0.027842231484017	0.000000000001316
Skanska Ab-B Shs	0.121957848612145	0.022440231460280
Skf Ab-B Shares	0.140934097318621	0.010792863661500
Electrolux Ab-Ser B	0.029703990159743	0.078533110523655
Kinnevik AB	0.100217568044991	0.106614848675072
Swedish Match Ab	0.000000000000000	0.021704009871380
Autoliv Inc-Swed Dep Receipt	0.000009179484781	0.038643710333476
Boliden Ab	0.031897784850201	0.000000010058815
Svenska Cellulosa Ab Sca-B	0.097296972455112	0.066813176198704
Securitas Ab-B Shs	0.000000000000000	0.000000002029802
Getinge Ab-B Shs	-0.000000000000000	0.000000046123691
Ssab Ab-A Shares	0.025776249883996	0.000004152100618

OMX - scenario 2		
	PSO	FWA
Astrazeneca Plc	-0.0000000000000000	0.007552635336410
Abb Ltd-Reg	0.0000000000000000	0.000867204615354
Atlas Copco Ab-B Shs	0.0000000000000000	0.000057900403032
Atlas Copco Ab-A Shs	0.0000000000000000	0.000001054661350
Investor Ab-B Shs	-0.0000000000000000	0.000001551785640
Hennes & Mauritz Ab-B Shs	0.127552298603559	0.000001764073911
Volvo Ab-B Shs	-0.0000000000000000	0.001849742432450
Nordea Bank Abp	0.060383483214403	0.017348468685130
Ericsson Lm-B Shs	0.0000000000000000	0.132086570984087
Assa Abloy Ab-B	0.0000000000000000	0.000010790419514
Skandinaviska Enskilda Ban-A	0.0000000000000000	0.417492460470016
Sandvik Ab	0.021445107090677	0.045706663460444
Telia Co Ab	-0.0000000000000000	0.000000014045564
Svenska Handelsbanken-A Shs	0.020180266822718	0.000000157008032
Hexagon Ab-B Shs	0.036008926295599	0.001404406585167
Swedbank Ab-A Shares	0.120615482365600	0.000009363321160
Tele2 Ab-B Shs	-0.0000000000000000	0.000000271422246
Alfa Laval Ab	0.114651233996600	0.000010811849200
Skanska Ab-B Shs	0.048918526579266	0.171972812483087
Skf Ab-B Shares	0.0000000000000000	0.000010444763127
Electrolux Ab-Ser B	0.147859303453626	0.000565369488393
Kinnevik AB	0.114692548810387	0.081607620510110
Swedish Match Ab	0.097321776875916	0.000000150632616
Autoliv Inc-Swed Dep Receipt	-0.0000000000000000	0.001363995940624
Boliden Ab	0.0000000000000000	0.000007752629899
Svenska Cellulosa Ab Sca-B	-0.0000000000000000	0.279341370711312
Securitas Ab-B Shs	0.0000000000000000	0.000022927692310
Getinge Ab-B Shs	0.090371918514339	0.000000014874886
Ssab Ab-A Shares	-0.0000000000000000	0.000002350573854

OMX - scenario 3		
	PSO	FWA
Astrazeneca Plc	0.078117292661872	0.000000191335042
Abb Ltd-Reg	0.040528488840409	0.004627778426275
Atlas Copco Ab-B Shs	0.0000000000000000	0.019994246464132
Atlas Copco Ab-A Shs	-0.0000000000000000	0.000130114228238
Investor Ab-B Shs	0.024740448692368	0.147597854544756
Hennes & Mauritz Ab-B Shs	-0.0000000000000000	0.000010004079785
Volvo Ab-B Shs	0.023772127301871	0.000012043670856
Nordea Bank Abp	0.092679184429408	0.031841455172438

Ericsson Lm-B Shs	-0.0000000000000000	0.021847618325391
Assa Abloy Ab-B	0.024429195784137	0.000972705512646
Skandinaviska Enskilda Ban-A	-0.0000000000000000	0.070180197903196
Sandvik Ab	0.0000000000000000	0.000113635240665
Telia Co Ab	0.021822232222089	0.017649438807691
Svenska Handelsbanken-A Shs	-0.0000000000000000	0.016900547466228
Hexagon Ab-B Shs	-0.000000000006850	0.131690592991945
Swedbank Ab-A Shares	0.137815339298431	0.000015585146965
Tele2 Ab-B Shs	0.041684642131251	0.000454479301222
Alfa Laval Ab	0.146273327157645	0.002882301980011
Skanska Ab-B Shs	0.044642906667146	0.000040816205080
Skf Ab-B Shares	0.0000000000000000	0.000004395782862
Electrolux Ab-Ser B	-0.0000000000000000	0.000122201283168
Kinnevik AB	0.022354186854387	0.069635876829754
Swedish Match Ab	-0.0000000000000000	0.000616447505222
Autoliv Inc-Swed Dep Receipt	0.096532308299113	0.051552755874985
Boliden Ab	0.145999623793728	0.242089342096004
Svenska Cellulosa Ab Sca-B	0.0000000000000000	0.000017764052570
Securitas Ab-B Shs	0.058721201264111	0.000044259847488
Getinge Ab-B Shs	0.0000000000000000	0.120966085934330
Ssab Ab-A Shares	0.0000000000000000	0.000001250033655

Bibliografia

Acerbi, C., Tasche, D. (2002), *Expected Shortfall: A Natural Coherent Alternative to Value at Risk*, Economic Notes, Banca Monte dei Paschi di Siena S.p.A., Vol. 31, No. 2-2002, pp. 379-388.

Acerbi, C., Tasche, D. (2002), *On the Coherence of Expected Shortfall*, working paper.

Ahmadi-Javid, A. (2012a), *Entropic value-at-risk: A new coherent risk measure*, Journal of Optimization Theory and Applications, 155 (3), 1105-1123.

Artzner, P., Delbaen, F., Eber, J.M., Heath, D. (1999), *Coherent measures of risk*, Mathematical Finance, 9(3), 203-227.

Blum, C., Roli, A., *Metaheuristics in Combinatorial Optimization: Overview and Conceptual Comparison*, ACM Computing Survey, 2003.

Cont, R. (2001), *Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues*, Quantitative Finance, 1, 223-236.

Corazza, M., corso di *Metodi per la gestione dei portafogli personali*, Università Ca' Foscari, Venezia, 2018.

Corazza, M., Fasano, G., Gusso, R., *Particle Swarm Optimization with non-smooth penalty reformulation, for a complex portfolio selection problem*, Applied Mathematics and Computation 224 (2013), 611-624.

Eberhart, R., Shi Y., *Particle swarm optimization: Development, applications and resources*, Proc. Congress on Evolutionary Computation 2001, Seoul, Korea. Piscataway, NJ: IEEE Service Centre.

Elton, E.J., Gruber, M.J., Brown, S.J., Goetzmann, W.N., *Teorie di portafoglio e analisi degli investimenti*, Apogeo, 2007.

Fister, I. Jr., Yang X., Fister I., Brest, J., Fister D. (2013), *A Brief Review of Nature-inspired Algorithms for Optimization*, *Elektrotehnicki Vestnik/Electrotechnical Review, English Edition*, 80(3), 1-7.

Gallocchio, C., 2017, *PSO vs FWA: due metaeuristiche per la selezione di portafogli basati sulle misure di rischio coerenti two-sided*, Università Ca' Foscari, Venezia.

Kennedy, J., Eberhart, R., *Particle Swarm Optimization*, Proc. IEEE International Conference on Neural Networks (Perth, Australia), IEEE Service Centre, Vol. 4, pp. 1942-1948, Piscataway, NJ, 1995.

Kennedy, J., Eberhart, R., *Swarm Intelligence*, Morgan Kaufmann of Academic Press, New York, 2001.

Linciano, N., Soccorso, P., *La rilevazione della tolleranza al rischio degli investitori attraverso il questionario*, Discussion Papers, CONSOB, 2012.

Luenberger, David G., *Investment Science*, Oxford University Press, Second Edition, 2013.

Lukasik, S., Kowalski, P.A., *Fully Informed Swarm Optimization Algorithms: Basic Concepts, Variants and Experimental Evaluation*, Proceedings of the 2014 Federated Conference on Computer Science and Information Systems, pp. 155-161, DOI: 10.15439/2014F377, ACSIS, Vol. 2.

Markowitz, H.M., *Portfolio Selection*, Journal of Finance, Vol. 7, No. 1, pp. 77-91, 1952.

Markowitz, H.M., *Portfolio selection: efficient diversification of investments*, New York, John Wiley & Sons, 1959.

Radovan, A., 2014, *La PSO per la selezione di portafogli basati sul Partitioned VaR e su vincoli misti-interi*, Università Ca' Foscari, Venezia.

Rizzo, M., 2018, *PSO e BFO per la gestione di hedge funds: il caso DeltaHedge*, Università Ca' Foscari, Venezia.

Sharpe, William F., *A Simplified Model for Portfolio Selection*, Management Science, Vol. 9, No. 2, University of Washington, January, 1963, pp. 277-293.

Sorensen, K., Sevaux, M., Glover, F., *A history of Metaheuristics*, in Mart, R., Pardalos, P., Resende, M., Handbook of Heuristics, Springer 2016.

Szegö, Giorgio P., *Portfolio Theory: With Application to Bank Asset Management*, New York, Academy Press, 1980 (pagina 8).

Tan, Y., Zhu, Y., *Fireworks Algorithm for Optimization*, Advances in Swarm Intelligence, Springer, 2010.

Von Neumann, J. & Morgenstern, O., *Theory of games and economic behaviour*, Princeton, New Jersey, US, Princeton University Press, 1947.

Zheng S., Li J., Janecek A., Tan Y., *A Cooperative Framework for Fireworks Algorithm*, IEEE/ACM Transactions on Computational Biology and Bioinformatics, 14(99), April 2015.

Sitografia

<http://bankpedia.org/index.php/it/133-italian/v/23638-value-at-risk-var-enciclopedia>

<http://econodinamica.blogspot.com/2015/10/cosa-e-entropia-economica.html>

<http://tateoblog.blogspot.com/p/lapproccio-euristico-e-le-tecniche.html>

<http://www.eticamente.net/2015/06/nicholas-georgescu-roegen-la-bioeconomia-e-la-legge-dellentropia.html>

<http://www.sapere.it/sapere/strumenti/studiafacile/fisica/Calore-e-termodinamica/La-termodinamica/L-entropia.html>

<http://www.treccani.it/enciclopedia/algoritmo/>

http://www.treccani.it/enciclopedia/bioeconomia_%28Dizionario-di-Economia-e-Finanza%29/

<https://risklover.wordpress.com/2009/05/16/misure-di-rischio-coerenti/>

<https://www.britannica.com/science/algorithm>

<https://www.britannica.com/science/NP-complete-problem>

<https://www.britannica.com/topic/computational-complexity>

<https://www.garzantilinguistica.it/ricerca/?q=algoritmo>

<https://www.geeksforgeeks.org/np-completeness-set-1/>