



Università  
Ca' Foscari  
Venezia

Corso di Laurea Magistrale

in Economia e Finanza

(ordinamento ex D.M. 270/2004)

Tesi di Laurea

# **PSO vs FWA: due metaeuristiche per la selezione di portafogli basati sulle misure di rischio coerenti two-sided**

**Relatore**

Ch. Prof. Marco Corazza

**Laureando**

Cristina Gallocchio

Matricola 843016

**Anno Accademico**

2016/ 2017



# Indice

Introduzione .....	5
LA TEORIA DELLA SELEZIONE DEL PORTAFOGLIO ALLA MARKOWITZ .....	7
1.1    Ipotesi del modello di Markowitz.....	7
1.2    Le due misure di sintesi: rendimento atteso e varianza .....	9
1.3    Il criterio media-varianza e il problema di selezione del portafoglio.....	11
1.4    Teoria dell'utilità attesa .....	15
1.5    Curve d'indifferenza e portafoglio ottimo.....	18
1.6    La diversificazione .....	20
FIREWORKS ALGORITHM.....	24
2.1    Dalle euristiche alle metaeuristiche .....	25
2.1.1    Le metaeuristiche .....	28
2.2    Swarm Intelligence .....	30
2.3    PSO-Particle Swarm Optimization.....	30
2.4    Le modifiche alla PSO.....	32
2.5    Topologia della popolazione .....	37
2.6    FWA-Fireworks Algorithm .....	39
2.7    Il funzionamento del FWA .....	40
2.7.1    Explosion Strenght.....	41
2.7.2    Explosion amplitude .....	42
2.7.3    Displacement operation .....	44
2.7.4    Mutation Operator.....	45
2.7.5    Mapping Rule.....	46
2.7.6    Selection strategy .....	47
2.7.7    Pseudo-codice FWA .....	48
2.8    Caratteristiche del Fireworks Algorithm .....	50
MISURA DI RISCHIO COERENTE TWO-SIDED.....	54
3.1    Critiche al modello media-varianza.....	55
3.2    Misure di rischio coerenti.....	58
3.2.1    Two-sided coherent risk measure .....	60
3.2.2    Costruzione della nuova misura di rischio.....	61

3.3	Applicazione della misura di rischio .....	62
3.4	Vincoli .....	63
3.5	Riformulazione del problema di selezione .....	66
3.5.1	Metodo delle penalità esatte.....	66
APPLICAZIONE DELLE METAEURISTICHE.....		69
4.1	Informazioni preliminari.....	69
4.2	Parametri.....	71
4.2.1	PSO-impostazione dei parametri .....	72
4.2.2	FWA-impostazione dei parametri.....	73
4.3	Applicazione.....	73
4.3.1	Scenari.....	76
Conclusione.....		84
Appendice A.....		85
Appendice B.....		91
Bibliografia.....		108

# Introduzione

La selezione dei portafogli in cui investire è uno dei temi cardine della finanza. Il problema di selezione del portafoglio è un problema di minimizzazione della misura di rischio che tiene conto di diversi vincoli. Il problema di selezione del portafoglio è stato introdotto da Markowitz nel 1952. Il modello in questione è stato il primo modello di selezione di portafoglio ad affrontare il problema di come investire in modo efficiente un dato capitale. Il modello di Markowitz, conosciuto anche con il nome di Modern Portfolio Theory, ha rivoluzionato il modo di fare investimenti nei titoli finanziari. Tuttavia il modello di Markowitz presenta alcuni limiti e un sistema d'ipotesi alquanto irrealistico. Queste limitazioni sono state riviste nel corso degli anni con lo scopo di rendere il problema di selezione del portafoglio più simile alle reali condizioni dei mercati finanziari e al comportamento degli investitori. Partendo dalle critiche al criterio utilizzato da Markowitz, ossia il criterio media-varianza, in questo elaborato s'intende proporre un nuovo modello di selezione di portafogli. La prima modifica al modello di selezione riguarda la misura di rischio utilizzata. L'impiego della varianza come misura di rischio nel portafoglio alla MARKOWITZ, misura la dispersione di una variabile casuale, non tenendo conto dei valori sopra la media e di quelli sotto la media (i cosiddetti upside e downside risk), con la conseguenza negativa di equiparare le perdite ai guadagni. Questa caratteristica della varianza è alquanto irrealistica per un investitore avverso al rischio

Nel modello di selezione del portafoglio proposto la varianza è sostituita da una misura di rischio coerente two-sided, accogliendo la tesi di Chen e Wang, secondo la quale le misure di rischio coerenti two-sided, non solo sono misure di rischio coerenti, ma a differenza di quest'ultime considerano tutta l'informazione della distribuzione dei rendimenti dei titoli. Le misure di rischio coerenti introdotte da Artzner et al (1999) sono misure di rischio che rispettano i quattro assiomi di coerenza. Il rispetto di tali assiomi implica una forma di congruenza tra l'incertezza avvertita dall'investitore e il numero che sintetizza la misura di rischio.

La seconda modifica al portafoglio alla Markowitz riguarda i vincoli. Il problema originale di selezione del portafoglio prevede infatti solo due vincoli: il vincolo di bilancio e il vincolo del rendimento desiderato. Il modello che si propone in questo elaborato prevede, oltre ai due vincoli appena citati, l'introduzione di altri sei vincoli, detti vincoli a variabili misto-interi. Questi vincoli hanno lo scopo di rendere il portafoglio simile alle reali condizioni d'investimento, ad esempio tenendo in considerazione i costi di transizione.

Il problema così riformulato è un problema di minimizzazione del rischio vincolato. La ricerca di soluzioni esatte per questo tipo di problema non è una soluzione fattibile in tempi ragionevoli. Per ovviare al problema si propone l'utilizzo di due diverse metaeuristiche, ossia degli algoritmi approssimativi che ricercano la soluzione ottima nell'area più promettente dello spazio delle soluzioni. Nello specifico si ricorre alla Particle Swarm Optimization (PSO) e al Fireworks Algorithm (FWA). Essendo entrambe le metaeuristiche nate come algoritmi risolutivi per problemi non vincolati si ricorre al metodo delle penalità esatte per riscrivere il problema vincolato in un problema non vincolato equivalente.

Infine sono condotti degli studi sulla capacità del modello proposto con entrambe le metaeuristiche attraverso degli appositi codici creati in Matlab

# Capitolo 1

## LA TEORIA DELLA SELEZIONE DEL PORTAFOGLIO ALLA MARKOWITZ

*Obiettivo del primo capitolo è introdurre la teoria della selezione del portafoglio alla Markowitz.*

*Dopo aver elencato le ipotesi su cui si basa il modello, si analizzano le due misure di sintesi della distribuzione dei rendimenti, ossia il rendimento atteso e la varianza, su cui si basa il criterio di scelta media-varianza. Successivamente si illustra come costruire la frontiera efficiente per un portafoglio formato da  $N$  titoli rischiosi, ponendo l'accento sulla correlazione tra gli asset. Infine si identifica il portafoglio ottimo tra quelli che costituiscono la frontiera in base al livello di propensione al rischio dell'investitore, ricorrendo ai concetti di funzione di utilità quadratica e curva d'indifferenza.*

### 1.1 Ipotesi del modello di Markowitz

Quando l'individuo percepisce un reddito superiore a quello che consuma, può decidere d'investire il surplus in due o più attività. Queste attività, detenute in opportune quantità, costituiscono il portafoglio dell'individuo. La composizione di tale portafoglio può essere frutto di decisioni casuali, ma può anche essere dovuta a scelte ben precise. In questo capitolo e nei successivi, l'attenzione è limitata solo alle attività finanziarie. Il portafoglio finanziario può essere definito come un vettore di dimensioni  $(1,N)$ , dove ogni elemento rappresenta la percentuale di ricchezza investita in quella determinata attività finanziaria:

$$x' = (x_1, \dots, x_n). \quad (1.1)$$

Uno dei primi modelli per determinare la composizione del portafoglio finanziario è sicuramente il modello di Markowitz (1952). Prima di illustrare tale modello, si elencano le ipotesi su cui si basano i modelli di selezione del portafoglio che si andranno a trattare in questo capitolo e nei successivi:

- l'orizzonte temporale è uniperiodale. Il modello non si pone il problema nel caso di reinvestimento della ricchezza all'epoca successiva. Quindi l'individuo seleziona il portafoglio nel periodo  $t$ , e questo verrà detenuto fino al periodo  $t + \Delta t$ , senza contemplare il caso in cui l'individuo disinvesta all'epoca  $t + \Delta t$  per poi reinvestire all'epoca  $t + 2\Delta t$  e poi fino all'epoca  $t + 3\Delta t$  e via dicendo;
- assioma di non sazietà. L'investitore posto davanti alla scelta tra due portafogli contenenti gli stessi asset, sceglierà il portafoglio che ha le quantità maggiori per almeno un bene. In altre parole l'individuo preferisce una quantità di ricchezza maggiore rispetto ad una minore.
- avversione al rischio. L'investitore, a parità di rendimento atteso, preferisce il portafoglio con il minor rischio;
- l'investitore si definisce razionale se rispetta le assunzioni di non sazietà di avversione al rischio.

Oltre alle ipotesi appena elencate, il modello di Markowitz richiede anche altre ipotesi:

- non sono presenti costi di transizione o tasse e tutti i beni sono perfettamente divisibili. Si ipotizza quindi l'assenza di frizioni nel mercato, scenario alquanto irrealistico;
- gli investitori sono price-taker, ossia non possono influenzare in nessun modo la distribuzione di probabilità dei rendimenti delle attività finanziarie; in altre parole nessuna loro scelta può cambiare i prezzi di mercato;
- le vendite allo scoperto<sup>1</sup> sono ammesse senza nessuna restrizione. È un'ipotesi forte, in quanto spesso nei periodi di crisi le vendite allo scoperto vengono vietate.

---

<sup>1</sup> La vendita allo scoperto, o in inglese short selling, consiste nella vendita di un titolo che ancora non si possiede. Il titolo viene prestato al venditore allo scoperto, spesso dietro al pagamento di commissioni. Nel momento in cui il venditore allo scoperto dovrà consegnare il titolo all'acquirente, sarà costretto a comprare l'attività finanziaria sul mercato, qualsiasi sia il suo prezzo. Ovviamente se il prezzo del titolo sarà sceso il venditore avrà un guadagno, viceversa una perdita. È facile intuire che questo tipo d'investitore scommette su un mercato al ribasso. Proprio

## 1.2 Le due misure di sintesi: rendimento atteso e varianza

In un ambiente rischioso, il risultato dell'investimento non è noto con certezza, bensì è rappresentato da una distribuzione di probabilità. Per esemplificare si pensi ad un investimento che fra un anno renda: il 20% se si verifica l'evento A, il 15% se si verifica l'evento B, e il 10% se si verifica l'evento C. I tre possibili eventi non sono altro che tre possibili stati del mondo, a cui viene associata una probabilità di realizzazione. Tuttavia, quando si ha a che fare con attività finanziarie, i possibili esiti non sono solamente tre, come nell'esempio. Di solito per sintetizzare l'informazione contenuta in una distribuzione vengono utilizzati i primi due momenti, ossia il valore atteso e la dispersione intorno al valore atteso: poiché si considera la distribuzione dei rendimenti di attività finanziarie, si parla, rispettivamente, di rendimento atteso e di rischio.

Il primo passo del modello di Markowitz consiste proprio nel misurare il rendimento e il rischio attraverso il rendimento atteso e la varianza<sup>2</sup>.

Se si indica con  $R$  una variabile casuale continua caratterizzata da una funzione di densità di probabilità  $f_R(\cdot)$ , allora il valore atteso sarà:

$$E(R) = \int_{-\infty}^{+\infty} t f(t) dt. \quad (1.2)$$

In generale, la media di una variabile casuale è un indice di posizione. Da un punto di vista finanziario si può considerare la media del tasso di rendimento di un investimento come una misura di redditività dell'investimento stesso.

Per quanto riguarda invece la varianza, in termini generale, indica un indice di variabilità. Da un punto di vista finanziario la varianza del tasso di rendimento di un investimento può essere vista come una misura di rischio dell'investimento stesso.

Se si indica con  $R$  una variabile casuale continua caratterizzata da una funzione di densità di probabilità  $f_R(\cdot)$ , allora la varianza è calcolata come:

$$Var(R) = \int_{-\infty}^{+\infty} (t - E(R))^2 f(t) dt \quad (1.3)$$

---

per questo motivo la vendita allo scoperto viene spesso vietata nei periodi di crisi, in quanto può favorire il crollo dei prezzi.

<sup>2</sup> Spesso è utilizzata come misura di rischio la radice quadrata della varianza, ossia la deviazione standard.

Descritte le misure calcolate per le singole attività si passa al calcolo del rendimento atteso e della varianza di portafoglio.

Indicando il tasso di rendimento di portafoglio come:

$$R_p = \sum_{i=1}^N x_i R_i \quad (1.4)$$

dove

$R_p$  è una variabile casuale,

$x_i$  con  $i = 1, \dots, N$  è il peso dell' $i$ -esimo titolo,

allora la media di  $R_p$  può essere calcolata come:

$$E(R_p) = \sum_{i=1}^N x_i r_i \quad (1.5)$$

o in notazione vettoriale,

$$r_p = x' r. \quad (1.6)$$

Analogamente la varianza di  $R_p$  può essere determinata come:

$$Var(R_p) = \sum_{i=1}^N x_i^2 \sigma_i^2 + 2 \sum_{i=1}^N \sum_{j=i+1}^N x_i x_j \sigma_{ij} \quad (1.7)$$

o in forma vettoriale

$$\sigma_p^2 = x' C x \quad (1.8)$$

dove  $C$  è la matrice varianza-covarianza.

Utilizzando il coefficiente di correlazione lineare di Bravais-Pearson, la varianza di portafoglio può essere riformulata come:

$$\sigma_p^2 = \sum_{i=1}^N x_i^2 \sigma_i^2 + 2 \sum_{i=1}^N \sum_{j=i+1}^N x_i x_j \rho_{ij} \sigma_i \sigma_j \quad (1.9)$$

### 1.3 Il criterio media-varianza e il problema di selezione del portafoglio

Descritti gli strumenti per misurare il rendimento futuro degli investimenti e la relativa misura di rischio, è necessario declinare un criterio per individuare i portafogli “migliori”. Il criterio utilizzato prende il nome di criterio media-varianza. Si tratta di un criterio efficiente, ossia un criterio che divide tutte le possibili scelte d’investimento in due insiemi mutualmente esclusivi: l’insieme efficiente e l’insieme inefficiente.

Se si considerano due variabili casuali  $R_1$  e  $R_2$ , allora  $R_1$  domina  $R_2$  in media-varianza se:

$$E(R_1) \geq E(R_2)$$

e

$$Var(R_1) \leq Var(R_2)$$

e

almeno una delle due disequaglianze è vera in senso stretto.

Per comprendere meglio il criterio media-varianza si propone il seguente esempio. Se si considera un titolo A con rendimento 25% e deviazione standard 10% e un titolo B con rendimento 10% e deviazione standard 40%, applicando il criterio media-varianza, si può concludere che il titolo A domina il titolo B. Nella Figura 1.1 è illustrata la situazione appena descritta.

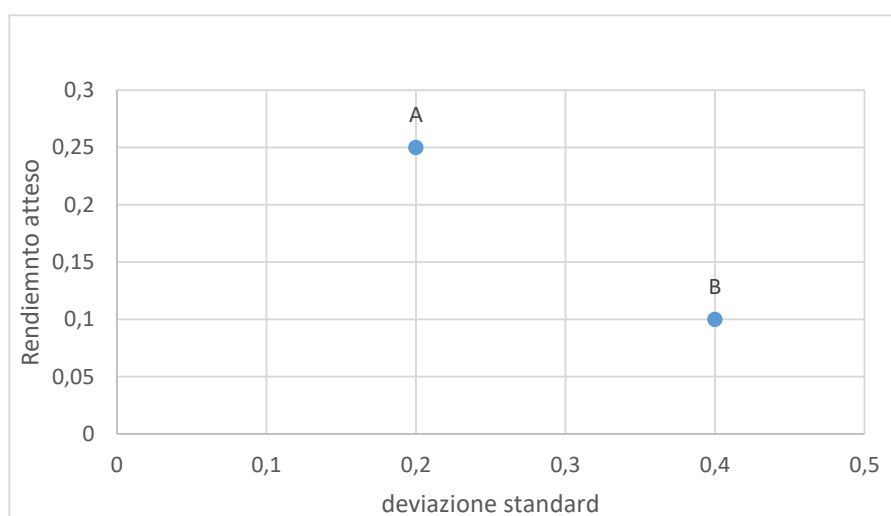


Figura 1.1: criterio media-varianza

Tuttavia non sempre il criterio media-varianza è in grado di ordinare un insieme di possibile scelte in cui investire. Può infatti accadere che ad esempio tra due scelte nessuna delle due domini l'altra in media-varianza. In questo caso entrambe le scelte d'investimento sono efficienti, e appartengono alla frontiera efficiente. La frontiera efficiente è quindi costituita dall'insieme dei portafogli efficienti, ossia quei portafogli che, fissato un certo rendimento desiderato, minimizzano il rischio di portafoglio, oppure analogamente, fissato un certo livello di rischio massimizzano il rendimento.

Se si considerano  $N$  titoli con rendimento variabile la determinazione della frontiera efficiente avviene risolvendo il seguente problema di minimizzazione:

$$\begin{aligned} & \min_{x_1, \dots, x_n} \sigma_p^2 & (1.10) \\ \text{s. t.} & \begin{cases} E(r_p) = \pi \\ x'e = 1 \end{cases} \end{aligned}$$

o in forma vettoriale:

$$\begin{aligned} & \min_{x_1, \dots, x_n} x'Cx & (1.11) \\ \text{s. t.} & \begin{cases} x'r = \pi \\ x'e = 1 \end{cases} \end{aligned}$$

dove

$\pi$  è il livello di rendimento atteso desiderato dall'investitore,

$x$  è il vettore colonna di ordine  $N$  contenente le quote d'investimento di ogni titolo,

$V$  è la matrice delle varianze e covarianze di ordine  $N$ ,

$r$  è il vettore di ordine  $N$  contenente i rendimenti di ciascun titolo,

$e$  è il vettore unitario di ordine  $N$ .

Il problema di ottimizzazione considerato da Markowitz richiede la minimizzazione del rischio fissato un certo livello di rendimento atteso desiderato dall'investitore. Il secondo vincolo, detto vincolo di bilancio, prevede che sia investito tutto il capitale disponibile.

La risoluzione del problema di ottimizzazione appena illustrato dà come risultato un vettore, in cui ogni elemento è il peso ottimo da investire in quella determinata attività finanziaria, formando così il portafoglio efficiente per il prescelto livello di rendimento atteso. Tuttavia perché esista un'unica soluzione è necessario che la matrice di varianza e covarianza  $C$  sia non-singolare<sup>3</sup> e definita positiva<sup>4</sup> e che almeno due titoli abbiano un rendimento atteso diverso, ossia  $R_i \neq R_j$  per qualche  $i, j = 1, \dots, N$ . Se e solo se queste condizioni sono verificate allora esiste un'unica soluzione. Per ottenere la formula del portafoglio ottimo bisogna risolvere il problema vincolato di minimizzazione della varianza di portafoglio. Per fare ciò si scrive dapprima la funzione lagrangiana:

$$L = x'Cx - \lambda_1(x'r - \pi) - \lambda_2(x'e - 1) \quad (1.12)$$

dove

$\lambda_1$  e  $\lambda_2$  sono i moltiplicatori di Lagrange.

Poi si pongono uguali a zero le derivate prime parziali di  $L$  e si mettono a sistema.

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

Con alcuni passaggi si arriva alla soluzione riportata qui sotto. Si specifica che il simbolo asterisco è usato per indicare il portafoglio efficiente.

$$x^* = \frac{(\gamma C^{-1}\mu - \beta C^{-1}e)\pi + (\alpha C^{-1}e - \beta C^{-1}\mu)}{\alpha\gamma - \beta^2} \quad (1.13)$$

dove

<sup>3</sup> Una matrice singolare è una matrice quadrata il cui determinante è uguale a zero; di conseguenza una matrice singolare è non invertibile.

<sup>4</sup> La matrice  $C$  è definita positiva se per ogni vettore  $M$  non nullo  $x'Cx > 0$

$$\alpha = \mu' C^{-1} \mu;$$

$$\beta = \mu' C^{-1} e = e' C^{-1} \mu;$$

$$\gamma = e' C^{-1} e;$$

Si osserva inoltre che:

- $E(R_{P^*}) = \bar{R}_{P^*} = x^{*\prime} \mu = \pi$  per il primo vincolo;
- $Var(R_{P^*}) = \sigma_{P^*}^2 = x^{*\prime} C x^* = \frac{\gamma \pi^2 - 2\beta \pi + \alpha}{\alpha \gamma - \beta^2}$ . L'equazione è quella di una parabola nel piano media-varianza;
- $StDev(R_{P^*}) = \sigma_{P^*} = (x^{*\prime} C x^*)^{\frac{1}{2}} = \left( \frac{\gamma \pi^2 - 2\beta \pi + \alpha}{\alpha \gamma - \beta^2} \right)^{\frac{1}{2}}$ . L'equazione descrive un'iperbole nel piano media-deviazione standard.

Nella Figura 1.9 sono rappresentate le combinazioni deviazione standard-rendimento atteso del portafoglio. Il codice Matlab 'Frontiera efficiente N>2', utilizzato per rappresentare la curva, è riportato in appendice A. In sintesi si sono considerati 10 titoli rischiosi con 30 osservazioni ciascuno. Le osservazioni sono state generate attraverso il comando RANDN che permette di campionare la distribuzione normale standardizzata. Si tratta di osservazioni generate casualmente e quindi non corrispondenti a nessuna serie storica di titoli esistenti: tuttavia, la distribuzione normale può rappresentare un andamento simile a quello di un titolo reale.

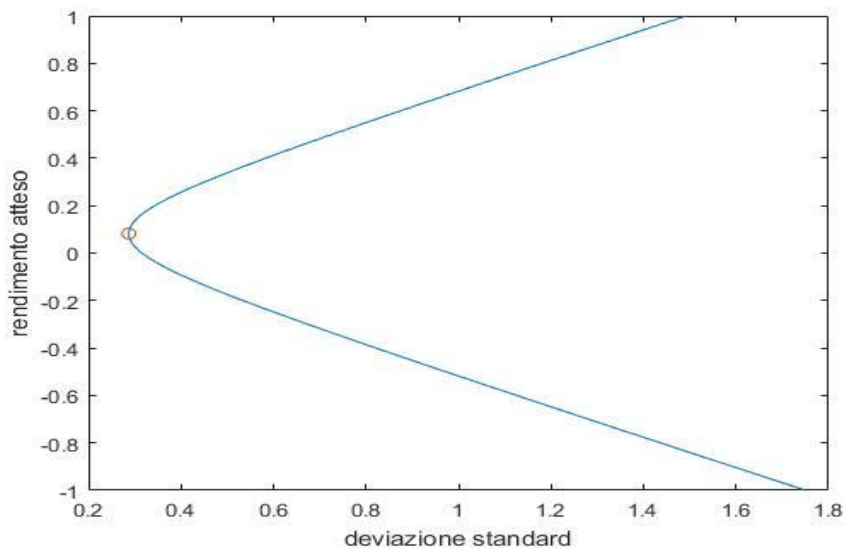


Figura 1.2: combinazioni rischio rendimento con  $N > 2$

Si è calcolata la matrice varianza-covarianza e si è fissato un vettore di rendimenti obiettivo, come richiede il primo vincolo del problema di minimizzazione; in questo esempio i rendimenti obiettivo sono 201. Il codice calcola il vettore composto da 10 pesi ottimi per ogni valore del vettore dei rendimenti obiettivo; si ottiene quindi una matrice di pesi ottimi di dimensioni (10, 201). Poi per ogni vettore colonna di pesi ottimi viene calcolata la deviazione standard del portafoglio corrispondente. Nel grafico è evidenziato il vertice dell'iperbole che ha le seguenti coordinate:

$$\left( \sigma_{p,v} = \left( \frac{1}{\gamma} \right)^{\frac{1}{2}}, r_{p,v} = \frac{\beta}{\gamma} \right) \quad (1.14)$$

I portafogli aventi un rendimento minore rispetto al rendimento del portafoglio corrispondente al vertice dell'iperbole sono dei portafogli inefficienti poiché sono dominati da quelli appartenenti alla frontiera efficiente, costituita da portafogli con rendimenti maggiori a parità di deviazione standard.

## 1.4 Teoria dell'utilità attesa

Finora si sono individuati una serie di portafogli efficienti. Tuttavia l'investitore dovrà determinare il portafoglio ottimo tra quelli efficienti nella quale investire. Per determinare il portafoglio ottimo bisogna prima definire un criterio per individuare tale portafoglio ottimo. Il criterio adottato è quello dell'utilità attesa, che permette di costruire una gerarchia delle preferenze dell'investitore, che tenga conta della personale avversione al rischio dell'investitore. Nello specifico, il criterio dell'utilità attesa permette di attribuire una scala soggettiva al valore monetario, ossia introdurre una funzione d'utilità  $u(x)$  del capitale  $x$ . In sostanza l'investitore adotterà un criterio di scelta tra due alternative aleatorie  $X_1$  e  $X_2$  basato sulla massimizzazione dell'utilità attesa. Di conseguenza l'investitore considererà un'operazione si scelta tra  $X_1$  e  $X_2$ :

- vantaggiosa se  $E[U(X_2)] - E[U(X_1)] > 0$
- indifferente se  $E[U(X_2)] - E[U(X_1)] = 0$

- svantaggiosa se  $E[U(X_2)] - E[U(X_1)] < 0$

La funzione di utilità è strettamente crescente poiché lo scopo dell'investitore è massimizzare il profitto:

$$u(X_2) > u(X_1), \quad \forall X_2 > X_1$$

Altra proprietà fondamentale della funzione di utilità è la concavità. Questa proprietà afferma che ad un incremento uguale del capitale corrisponde un aumento dell'utilità tanto minore quanto maggiore è il capitale posseduto. Per chiarire questo concetto si pensi ad esempio ad un individuo A che possieda un capitale di 100 euro e ad un individuo B che invece possieda un capitale di 100000 euro. Se i due capitali aumentano entrambi di 3000 euro l'utilità che avrà l'individuo A sarà maggiore di quella dell'individuo B.

Queste due proprietà si ritrovano anche nella funzione di utilità utilizzata nel modello di Markowitz, ossia la funzione di utilità quadratica:

$$u(R_p) = R_p - \left(\frac{a}{2}\right) R_p^2 \text{ con } a \in R^+ \quad (1.27)$$

Il parametro  $a$  riflette l'avversione al rischio dell'investitore; maggiore è questo parametro e maggiore è l'avversione al rischio dell'investitore.

La funzione di utilità quadratica dipende solo da media e varianza. Se infatti si prende il valore atteso della funzione si ottiene:

$$E[U(R_p)] = E(R_p) - \left(\frac{1}{2}\right) a E(R_p^2) = \mu - \left(\frac{1}{2}\right) a (\sigma^2 + \mu^2) \quad (1.15)$$

poiché per definizione  $\sigma^2 = E(R_p^2) - [E(R_p)]^2$ . La scelta della funzione di utilità quadratica nel modello di Markowitz non è quindi casuale, infatti è consistente con il criterio media-varianza.

Ritornando alle due proprietà di crescita e concavità illustrate prima, si rappresenta nella Figura (1.3)<sup>5</sup> la funzione di utilità quadratica. In ordinata sono rappresentati i valori dell'utilità attesa, mentre in ascissa i valori della ricchezza.

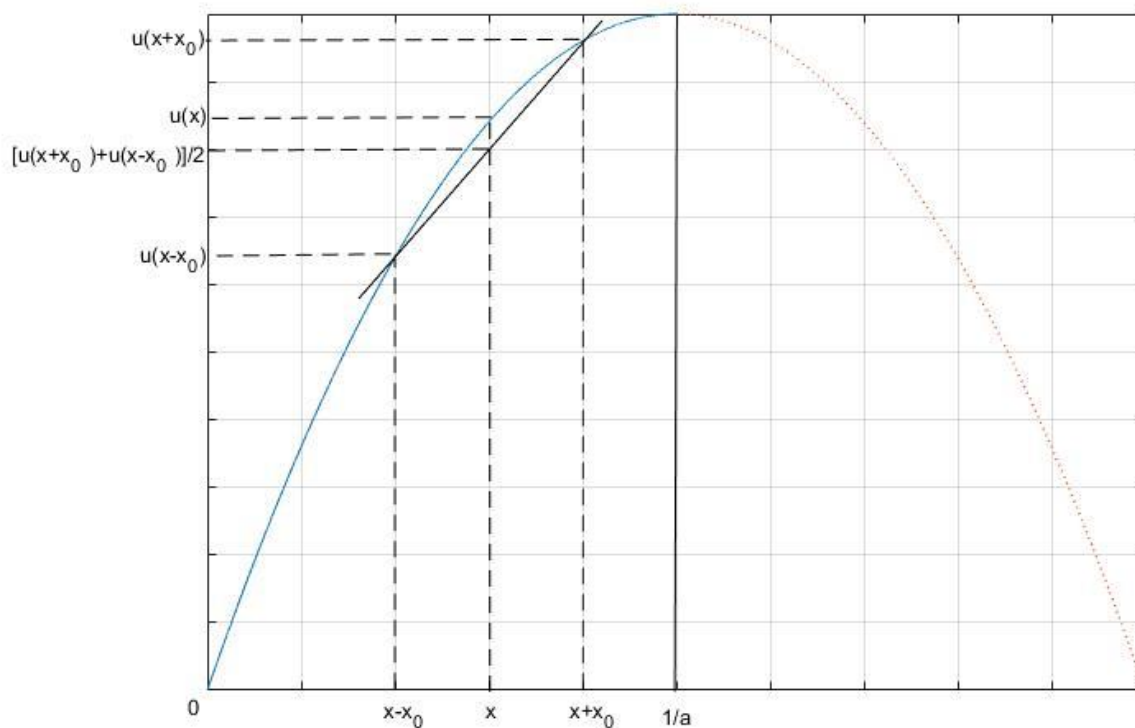


Figura 1.3: funzione di utilità quadratica

La funzione è strettamente crescente per le  $x$  comprese tra  $0$  e  $1/a$ ; il dominio viene quindi ridotto a questo intervallo per garantire la crescita della funzione. La concavità è assicurata dalla non negatività del parametro  $x$ , infatti all'incremento di capitale da  $x - x_0$  a  $x$  corrisponde un incremento dell'utilità maggiore rispetto all'incremento di capitale da  $x$  a  $x + x_0$ , quindi

$$u(x) - u(x - x_0) > u(x + x_0) - u(x)$$

Esplicitando la disuguaglianza rispetto a  $u(x)$ , si ottiene

$$u(x) > \frac{u(x + x_0) + u(x - x_0)}{2} \quad (1.16)$$

<sup>5</sup> la Figura 1.5 è stato ottenuto utilizzando il codice Matlab 'Funzione utilità quadratica' riportato in Appendice B.

La disuguaglianza (1.16) indica che l'investitore è avverso al rischio: infatti se un individuo con un capitale iniziale pari a  $x$ , scommettesse una quantità  $x_0$ , la media delle due utilità che si avrebbe in caso di vincita e perdita, sarebbe minore dell'utilità che l'investitore avrebbe astenendosi dal scommettere.

Più in generale, il comportamento dell'individuo nei confronti del rischio può essere dedotto dalla derivata seconda della funzione di utilità, infatti:

- se  $u''(x) < 0$  allora l'individuo è avverso al rischio
- se  $u''(x) = 0$  allora l'individuo è indifferente al rischio
- se  $u''(x) > 0$  allora l'individuo è propenso al rischio

La derivata prima della funzione di utilità invece individua l'utilità marginale di  $x$  ed è positiva su tutto il dominio poiché la funzione d'utilità è strettamente crescente.

## 1.5 Curve d'indifferenza e portafoglio ottimo

In questo paragrafo si deriva dalla curva d'utilità quadratica, l'equazione della curva d'indifferenza. La curva d'indifferenza non è altro che la rappresentazione sul piano cartesiano delle scelte di consumo che danno al consumatore la stessa utilità. Nel caso del problema di selezione del portafoglio ottimo si parlerà delle scelte d'investimento che danno all'investitore lo stesso grado d'utilità.

Nello specifico le curve d'indifferenza permettono d'individuare le combinazioni rischio-rendimento considerate equivalenti dall'investitore. La curva d'indifferenza è costruita nello spazio  $(\sigma, \mu)$  ed è formata da tutte le combinazioni rischio-rendimento che hanno la stessa utilità attesa per l'investitore e che quindi sono indifferenti. La curva d'indifferenza può essere ricavata esplicitando la deviazione standard dal valore atteso della funzione d'utilità quadratica, per un fissato valore di utilità attesa :  $\bar{U}$

$$\bar{U} = \mu - \left(\frac{1}{2}\right) a(\sigma^2 + \mu^2)$$

$$\sigma^2 = -\frac{2}{a} * \bar{U} + 2 * \frac{\mu}{a} - \mu^2$$

$$\sigma = \sqrt{-\frac{2}{a} * \bar{U} + 2 * \frac{\mu}{a} - \mu^2} \quad (1.17)$$

Per capire l'andamento delle curve d'indifferenza, si deriva:

$$d[E(U)] = 0 = d\mu - a\mu d\mu - a\sigma d\sigma$$

$$\frac{d\mu}{d\sigma} = \frac{a\sigma}{1 - a\mu} > 0$$

Essendo la derivata prima, rispetto a  $\sigma$ , positiva, si può affermare che le curve d'indifferenza sono crescenti, ossia quando cresce  $\sigma$  aumenta anche  $\mu$ , ciò per mantenere l'utilità attesa costante.

Calcolando la derivata seconda si ottiene:

$$\frac{d^2\mu}{d\sigma^2} = \frac{a(1 - a\mu) + a^2\sigma \left(\frac{d\mu}{d\sigma}\right)}{(1 - a\mu)^2} > 0$$

Poiché la derivata seconda è positiva, la curva d'indifferenza è convessa, ossia all'aumentare della deviazione standard  $\sigma$ , il rendimento atteso  $\mu$  cresce più che proporzionalmente.

Analizzando la derivata prima si può notare che bisogna porre un vincolo di non-negatività al denominatore perché la curva sia crescente e convessa. Infatti solo per valori di  $\mu < \frac{1}{a}$  le considerazioni fatte risultano vere.

Nella Figura 1.10 sono state rappresentate alcune curve di indifferenza per diversi livelli di utilità attesa. Si è utilizzato il codice Matlab 'Curve d'indifferenza' riportato in appendice B; in sintesi il codice crea un vettore di possibili rendimenti attesi, restringendo il range per valori di  $\mu < \frac{1}{a}$ . Attraverso l'equazione (1.29) si calcola la corrispondente varianza per ogni valore di  $\mu$ . Come si può vedere dal grafico sono stati calcolati diversi livelli di utilità attesa, in particolare come indicato nella leggenda per  $\bar{U} = 1, 2, 3, 4$ . All'aumentare dell'utilità, a parità di rendimento atteso, diminuisce la deviazione standard. Un'analogia considerazione vale per l'aumento del rendimento atteso: all'aumentare dell'utilità, il rendimento atteso aumenta a parità di deviazione standard.

L'individuo preferirà collocarsi sulla curva d'indifferenza più "in alto", poiché il livello di utilità è maggiore. Tuttavia il portafoglio ottimo dovrà essere identificato all'interno dei portafogli efficienti. Si deduce, in conclusione, che in generale il portafoglio ottimo corrisponda al punto di tangenza tra la frontiera efficiente e la curva d'indifferenza con l'utilità maggiore possibile.

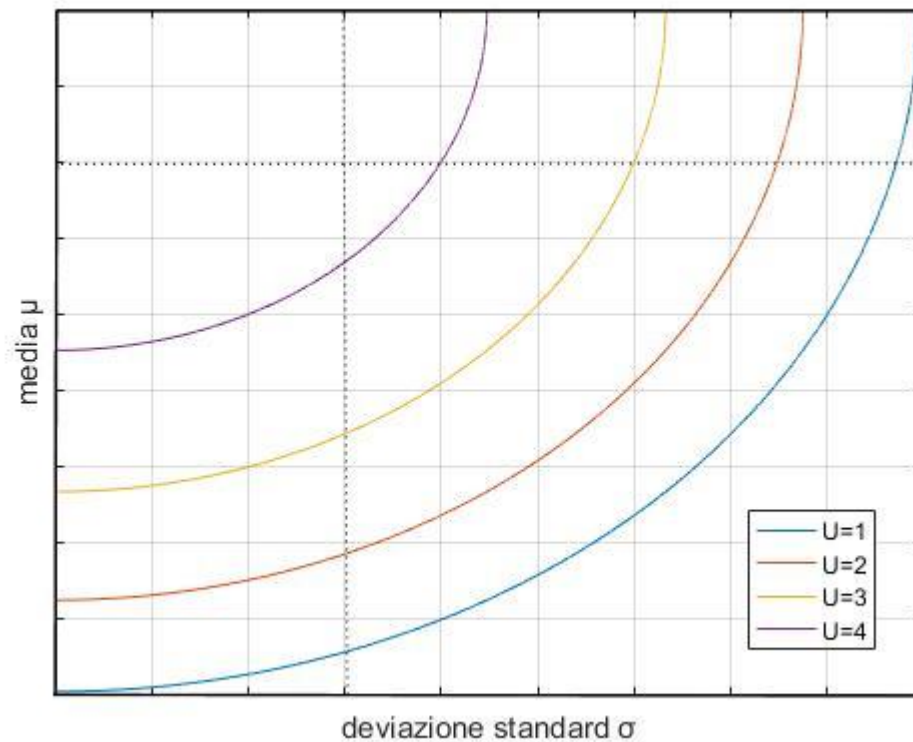


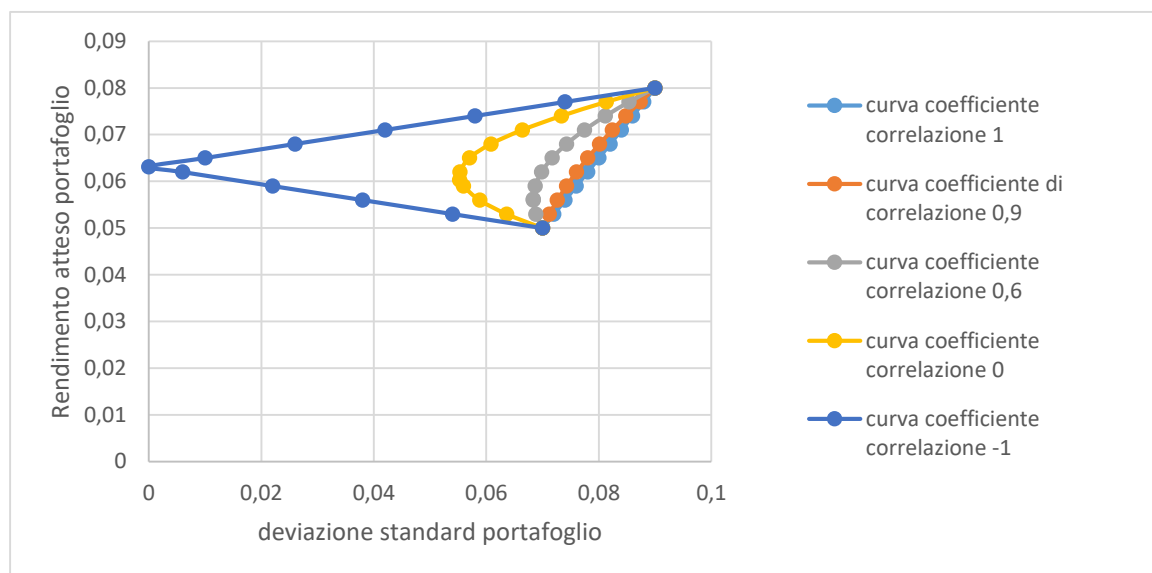
Figura 1.4: Curve d'indifferenza

## 1.6 La diversificazione

Uno degli aspetti più innovativi del modello di Markowitz è sicuramente il concetto di diversificazione nel mondo finanziario. Si può usufruire del beneficio della diversificazione del portafoglio nel momento in cui investendo in due o più attività si può sfruttare il coefficiente di correlazione, per ridurre il rischio del portafoglio ad un livello inferiore a quello del titolo meno rischioso contenuto nel portafoglio.

Per comprendere meglio il concetto si considera la rappresentazione in Figura 1.8: le diverse combinazioni rischio-rendimento sono state ottenute variando il coefficiente di correlazione. Per semplicità le curve sono state costruite con un portafoglio di soli due titoli rischiosi.

Figura 1.5: confronto 5 diversi coefficienti di correlazione



Analizzando le diverse curve si possono fare le seguenti osservazioni:

- la curva con coefficiente di correlazione pari a 1 è costituita interamente da portafogli efficienti. Solo in questo caso la curva delle combinazioni rischio-rendimento coincide con la curva della frontiera efficiente, mentre per valori del coefficiente di correlazione diversi da 1 si può individuare un tratto di curva formato da portafogli efficienti, e un altro tratto formato invece da portafogli inefficienti. Ad esempio si riporta il grafico, in Figura 1.6, della combinazione rischio-rendimento tra due titoli rischiosi A e B con  $\rho_{AB} = 0,6$ . La curva AB rappresenta le combinazioni rendimento-deviazione standard; il tratto di curva AM, con M escluso, rappresenta i portafogli non efficienti, con M portafoglio di minima-varianza. Il portafoglio di minima varianza è il portafoglio con il rischio minimo tra le combinazioni rischio-rendimento della curva. Il tratto di curva MB, con M incluso, rappresenta la frontiera efficiente formata dai portafogli efficienti.

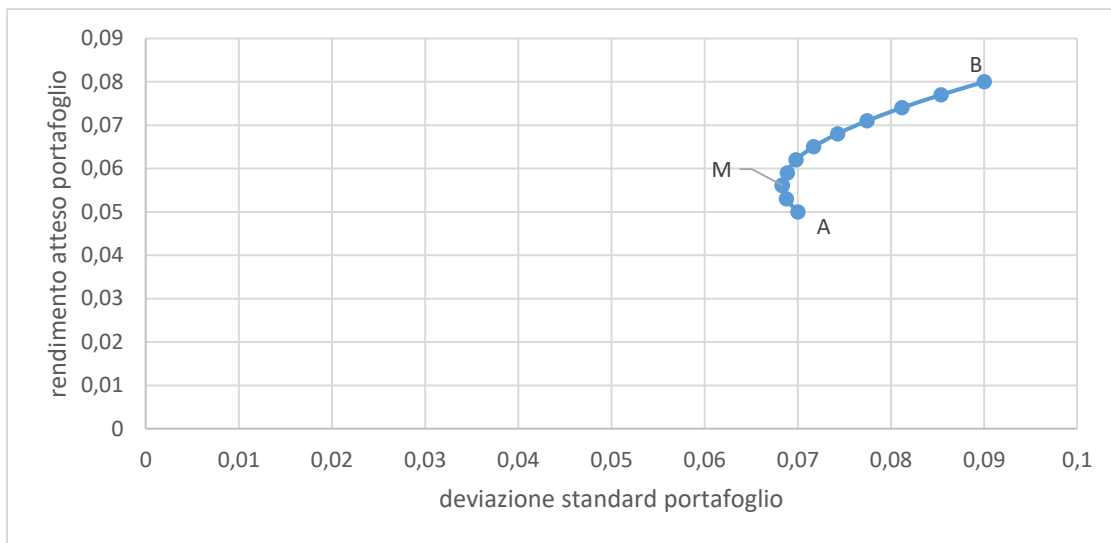


Figura 1.6: combinazioni rischio-rendimento tra due titoli rischio con  $\rho_{AB} = 0,6$

Per valori del coefficiente di correlazione appena inferiori a 1 si possono notare i primi benefici della diversificazione. Ad esempio nella Figura 1.7, si può vedere come i portafogli formati dai due titoli con coefficiente di correlazione pari a 0,9 dominano i portafogli con coefficiente di correlazione pari a 1, poiché a parità di rendimento la deviazione standard è minore.

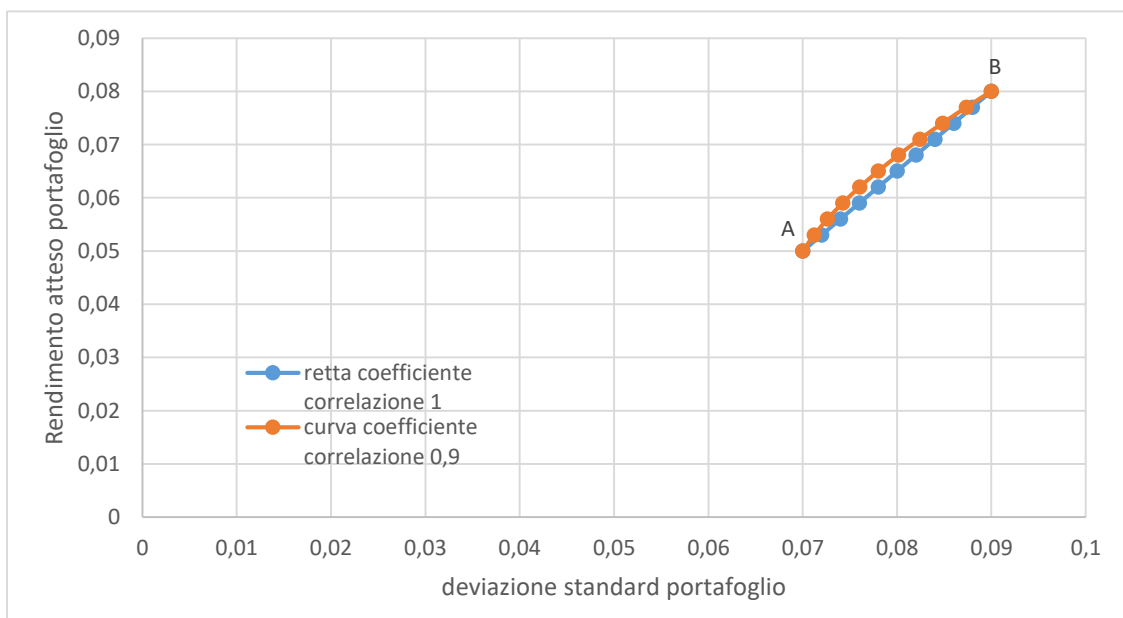
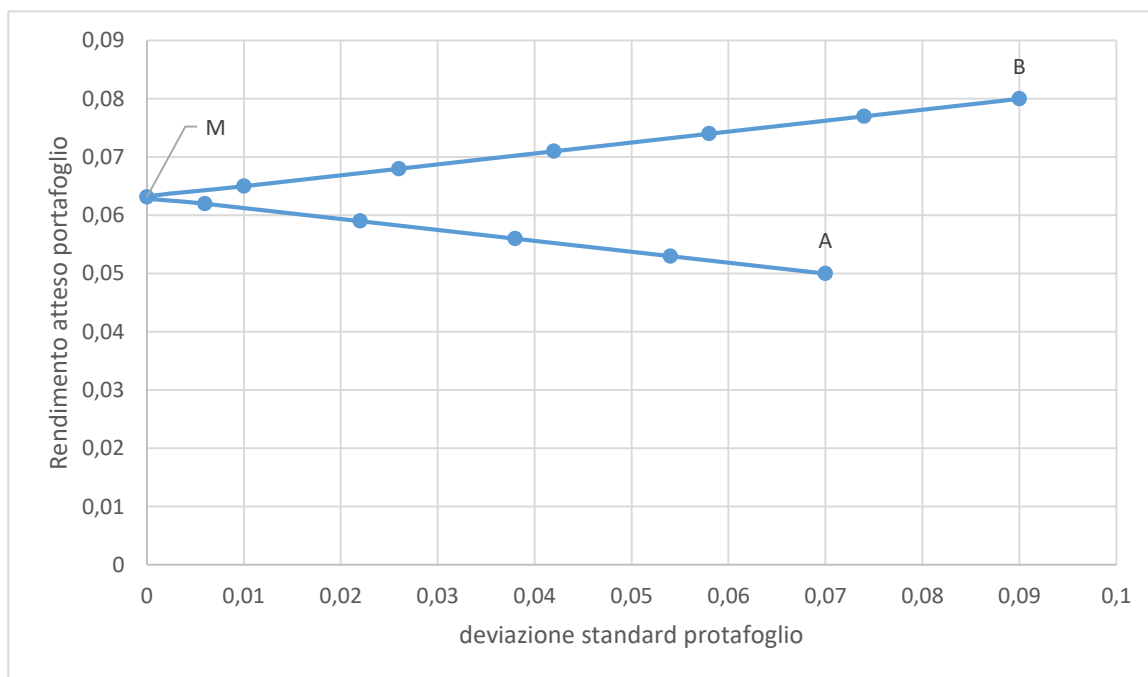


Figura 1.7: confronto curva con  $\rho_{AB} = 1$  e curva con  $\rho_{AB} = 0,9$

- il beneficio della diversificazione è massimo per  $\text{con } \rho_{AB} = -1$ .
- Nella Figura 1.8 si può vedere come gli effetti della diversificazione, quando il coefficiente di correlazione è perfettamente negativo, siano massimi; infatti dalla combinazione dei due titoli è possibile ottenere un portafoglio con varianza nulla.

Figura 1.8: combinazioni rischio-rendimento tra due titoli rischio con  $\rho_{AB} = -1$



## Capitolo 2

# PARTICLE SWARM OPTIMIZATION

## E

### FIREWORKS ALGORITHM

*In questo capitolo sono presentate due metaeuristiche per la risoluzione dei problemi di ottimizzazione: la Particle Swarm Optimization (PSO) e il Firework Algorithm (FWA).*

*Dopo aver fornito una panoramica generale del concetto di metaeuristica e di Swarm Intelligence, si analizza nel dettaglio la PSO, soffermandosi sull'algoritmo originale introdotto da Kennedy e Eberhart. Successivamente si propongono alcune modifiche all'algoritmo di base, come ad esempio l'introduzione del peso d'inerzia. Infine vengono analizzate le diverse topologie che possono essere utilizzate nella PSO.*

*Per quanto riguarda il FWA, il capitolo prevede l'analisi dell'algoritmo, dei quattro componenti chiave, e delle principali caratteristiche che lo contraddistinguono.*

*In conclusione vengono confrontate le due metaeuristiche analizzate, individuando sia i punti in comune che le differenze.*

## 2.1 Dalle euristiche alle metaeuristiche

La Particle Swarm Optimization (PSO) e il Fireworks Algorithm (FWA) appartengono entrambe alla classe degli algoritmi<sup>6</sup> approssimati, e più precisamente alle metaeuristiche. Questi algoritmi permettono di risolvere problemi di Ottimizzazione Combinatoria (CO) a cui appartiene il problema di selezione del portafoglio ottimo. Nei problemi di ottimizzazione combinatoria si ricerca una soluzione in un insieme finito di elementi. Questa soluzione può essere ad esempio, un numero intero, o un sottoinsieme di numeri interi, o una permutazione.

Un problema di Ottimizzazione Combinatoria, definito come  $P = (S, f)$ , è formato da:

- un insieme di variabili  $X = (x_1, \dots, x_n)$
- i domini variabili  $D_1, \dots, D_n$ ;
- una funzione obiettivo  $f$  da minimizzare<sup>7</sup>, dove  $f: D_1 \times \dots \times D_n \rightarrow R^+$ .
- dei vincoli sulle variabili;

L'insieme di tutte le possibili combinazioni è dato da:

$$S = \{s = \{(x_1, v_1), \dots, (x_n, v_n)\} | v_i \in D_i, s \text{ soddisfa tutti i vincoli}\} \quad (2.1)$$

$S$  è comunemente chiamato spazio delle soluzioni (o di ricerca), poiché ogni elemento contenuto in esso può essere visto come la possibile soluzione.

Per risolvere il problema di CO bisogna trovare la soluzione ottima  $s^* \in S$  che minimizza la funzione obiettivo, in modo che

$$f(s^*) \leq f(s) \quad \forall s \in S \quad (2.2)$$

$s^*$  è la soluzione ottima globale di  $(S, f)$  e  $S^* \subseteq S$  è l'insieme delle soluzioni ottime globali.

Per risolvere i problemi di ottimizzazione combinatoria si può ricorrere a due tipologie di

---

<sup>6</sup> L'algoritmo o metodo è un insieme ordinato di operazioni che porta alla soluzione di una certa classe di problemi. L'algoritmo riceve dei dati in ingresso, detti input e produce dei dati in uscita, detti output.

<sup>7</sup> Nel caso di un problema di massimizzazione basterà minimizzare la funzione  $-f$ .

algoritmi: gli algoritmi esatti, detti anche completi, e gli algoritmi approssimativi, detti anche euristici.

Gli algoritmi completi trovano la soluzione ottima in un tempo limitato per problemi di dimensioni finite. Tuttavia per i problemi *NP-hard*, non esiste alcun algoritmo completo che riesca a trovare la soluzione ottima in tempo polinomiale. In nota si riportano i concetti di tempo polinomiale e di problemi *NP-hard*<sup>8</sup>. In altre parole, un algoritmo completo trova una soluzione a un problema *NP-hard* in tempi non ragionevoli.

Gli algoritmi euristici, a differenza degli algoritmi completi, permettono di trovare una ‘buona’ soluzione al problema, sacrificando la ricerca della soluzione ottima a favore di tempi computazionali ragionevoli. Sono quindi utilizzati per la risoluzione di problemi *NP-hard*.

Gli algoritmi euristici possono essere suddivisi in tre categorie:

- algoritmi costruttivi. Questi algoritmi sono orientati alla costruzione di una soluzione ammissibile e prevedono tre passi fondamentali:
  - l’inizializzazione, dove viene scelto un elemento di partenza per costruire la soluzione parziale;
  - la selezione di un nuovo elemento, attraverso un criterio prescelto, da aggiungere alla soluzione parziale;
  - il criterio di arresto; se la soluzione è ammissibile allora l’algoritmo si ferma, altrimenti si ritorna al punto precedente.

Gli algoritmi costruttivi sono algoritmi veloci, ma spesso la soluzione trovata è meno precisa rispetto agli algoritmi di ricerca locale;

---

<sup>8</sup> Per spiegare il concetto si ricorre brevemente alla definizione di complessità temporale: in informatica, la complessità temporale di un algoritmo misura il tempo che impiega l’algoritmo a essere eseguito in funzione della lunghezza della sequenza di caratteri che formano l’input. Ad esempio per ordinare tremila elementi ci vorrà un tempo maggiore rispetto a ordinarne dieci: per definire quanto tempo è necessario si utilizza l’ordine di grandezza, indicato con la notazione  $O(f(n))$ . Quindi, considerato un numero iniziale di input  $n$ , la complessità di un algoritmo è  $O(f(n))$  se il numero di operazioni per eseguire l’algoritmo è un multiplo di  $f(n)$ .

In particolare un algoritmo in tempo polinomiale è 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 qualsiasi input di dimensione  $n$ . I problemi per cui è conosciuto un algoritmo che necessita di un numero polinomiale di operazioni per risolverli, fanno parte della classe di complessità  $P$  (polynomial). Invece per problemi della classe  $NP$ , (non deterministic polynomial), non sono attualmente noti algoritmi di complessità polinomiale che riescano a risolverli, ma potrebbero essere risolti in un tempo polinomiale grazie a un computer non deterministico, chiamato macchina di Turing: non essendo negli obiettivi della tesi, si definisce, in modo informale, la macchina di Turing come una macchina ideale che trova sempre la scelta giusta e quindi riesce a risolvere un problema  $NP$  in un tempo polinomiale. Infine i problemi *NP-hard* possono essere definiti come una classe di problemi almeno difficile quanto i più difficili problemi delle classi  $P$  e  $NP$ .

- algoritmi di ricerca locale, anche detti migliorativi. Questi algoritmi prevedono quattro fasi fondamentali:
  - l’inizializzazione, dove viene scelta una soluzione iniziale per innescare il processo di ricerca;
  - l’individuazione del vicinato, che consente di individuare un intorno della soluzione corrente;
  - la scelta di una nuova soluzione, dove s’individua la soluzione migliore del vicinato. La struttura del vicinato è definita come una funzione  $N: S \rightarrow 2^S$  che assegna ad ogni  $s \in S$  un insieme di vicini  $N(s) \subseteq S$ , dove  $N(s)$  è il vicinato di  $s$ .

La soluzione minima locale può quindi essere definita come la soluzione per la quale la funzione obiettivo è minore rispetto alla funzione obiettivo di tutti gli altri elementi del vicinato. Quindi  $\hat{s}$  sarà un minimo locale rispetto alla struttura del vicinato se:

$$f(\hat{s}) \leq f(s), \quad \forall s \in N(s) \quad (2.4)$$

dove  $N(s)$  è il vicinato di  $s$ .

- il criterio di arresto: se la nuova soluzione individuata è migliore rispetto alla soluzione iniziale allora si torna al secondo punto, altrimenti l’algoritmo si ferma.
- le metaeuristiche. Il termine meta-euristica è formato da due parole di origine greca: euristica deriva dal verbo *heuriskein* che significa “cercare”, mentre il suffisso *meta* significa “oltre, ad un livello superiore”. Da questa definizione si può intuire come le metaeuristiche non sono delle semplici euristiche come le due precedenti, ma mirano a trovare una soluzione esplorando lo spazio di ricerca in modo efficiente ed efficace. Mentre gli algoritmi costruttivi individuano una soluzione ammissibile e gli algoritmi migliorativi partono da una soluzione ammissibile cercando di migliorala per giungere a un ottimo locale, le metaeuristiche migliorano la qualità delle soluzioni individuate con i metodi costruttivi e di ricerca locale.

Fanno parte delle metaeuristiche sia la Particle Swarm Optimization che il Fireworks Algorithm.

---

Nel paragrafo successivo sarà analizzata in dettaglio quest'ultima categoria di algoritmi approssimativi, le metaeuristiche, dandone una definizione e una classificazione.

### 2.1.1 Le metaeuristiche

Negli ultimi 30 anni le metaeuristiche hanno conosciuto uno sviluppo notevole, tanto che la stessa definizione di metaeuristica è cambiata nel corso degli anni. Sebbene siano state proposte molte definizioni, ancora oggi non vi è una definizione condivisa da tutti i ricercatori. In questo elaborato si è scelto di utilizzare una definizione recente di Sörensen e Glover del 2013.

“A metaheuristic is a high-level problem-independent algorithmic framework that provides 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 algorithms. The term is also used to refer 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 restituisce in tempi ragionevoli, una soluzione che è una buona approssimazione della soluzione ottima, la metaeuristica invece approfondisce la ricerca della soluzione ottima nelle zone più promettenti nello spazio delle soluzioni.

Sebbene lo studio scientifico delle metaeuristiche sia recente, il loro utilizzo da parte dell'uomo e degli animali risale alla preistoria e ha accompagnato la Terra nelle sue fasi evolutive. Per la mente umana è abbastanza intuitivo risolvere problemi di ottimizzazione sin dalla giovane età: infatti la capacità di risolvere problemi in modo veloce e corretto è uno dei fondamenti su cui si basa la sopravvivenza della specie. I problemi di ottimizzazione sono all'ordine del giorno: si pensi ad esempio alla scelta di quale strada percorrere per andare a lavoro cercando di scegliere la più breve, quella senza traffico e magari senza pedaggi, oppure al riempire la lavastoviglie cercando di ottimizzare lo spazio, fino ad arrivare al problema di ottimizzazione in questione, ossia quale portafoglio selezionare per massimizzare il profitto e minimizzare il rischio.

Le metaeuristiche sono sostanzialmente basate su due concetti: la diversificazione, che si riferisce all'esplorazione dello spazio di ricerca e l'intensificazione, che invece si riferisce allo sfruttamento della ricerca nello spazio fatta fino a quel momento. Queste due caratteristiche

devono essere bilanciate, da un lato per trovare velocemente zone promettenti nella ricerca della soluzione ottima, e dall'altro per evitare di perdere tempo in regioni già esplorate o non promettenti.

Di seguito si darà una breve panoramica delle metaeuristiche, classificandole in base a diverse caratteristiche:

- *nature inspired vs non-nature inspired*. Questa distinzione si basa sull'origine della meta-euristica; si distinguono, infatti, algoritmi come ad esempio l'Ant Algorithm, che s'ispirano al comportamento della natura, nel caso specificato alle colonie di formiche, e algoritmi che invece non s'ispirano al comportamento della natura, come ad esempio il Fireworks Algorithm, che s'ispira all'esplosione dei fuochi d'artificio. Questa classificazione non è molto utile nella pratica in quanto vi sono metodi ibridi che potrebbero appartenere a tutte e due le categorie, ma anche per le metaeuristiche non ibride non sempre è chiaro come classificarle;
- *population-based vs single point search*. La differenza tra le due categorie si basa sostanzialmente nel numero di soluzioni usate contemporaneamente. Gli algoritmi che lavorano su singole soluzioni sono detti trajectory method. Viceversa le metaeuristiche population based generano e gestiscono un insieme di soluzioni, chiamato appunto popolazione;
- *funzione obiettivo statica vs funzione obiettivo dinamica*. Mentre in alcune metaeuristiche la funzione obiettivo rimane la stessa, in altre viene modificata durante la ricerca per permettere di "scappare" da eventuali minimi locali cambiando zona;
- *una struttura del vicinato vs molteplici strutture del vicinato*. Sebbene molte metaeuristiche lavorino con una sola struttura del vicinato, ossia la forma della topologia dello spazio di ricerca non cambia durante l'algoritmo, alcune, invece, usano un insieme di strutture dei vicinati che permettono di diversificare la ricerca cambiando la forma della topologia dello spazio.
- *uso della memoria vs non uso della memoria*. Gli algoritmi che non utilizzano la memoria passata rappresentano un processo di Markov<sup>9</sup>, ossia per determinare l'azione successiva utilizzano solo l'informazione attuale. Invece tra gli algoritmi che usano la memoria si può distinguere tra quelli con memoria a breve termine e quelli con memoria a lungo termine. Mentre i primi si avvalgono solo d'informazioni recenti, i secondi

---

<sup>9</sup> Un processo stocastico markowiano 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.

utilizzano dei parametri di sintesi per poter muoversi anche in base alla memoria di lungo termine.

## 2.2 Swarm Intelligence

I due algoritmi che si andranno a confrontare, PSO e FWA, appartengono entrambi alla classe di algoritmi di tipo Swarm Intelligence (SI).

La SI è un ramo della Computazione Evolutiva<sup>10</sup>, che insieme agli altri metodi citati in nota 6, costituisce una famiglia di algoritmi per l'ottimizzazione globale che simula i fenomeni naturali e i processi biologici. In particolare la SI è considerata come un comportamento collettivo di un sistema decentralizzato e auto-organizzato; in altre parole, è il comportamento di molti individui che dimostrano un'intelligenza collettiva senza però avere una componente centrale che supervisiona l'intero processo. Ogni individuo della popolazione interagisce, direttamente o indirettamente, con gli individui dell'ambiente circostante. Sebbene questi individui seguano delle regole molto semplici, in realtà la comunicazione tra di essi porta ad un complesso comportamento globale, ben oltre le capacità di ogni singolo soggetto. Esempi in natura di SI sono i comportamenti degli stormi di uccelli, delle colonie di api, delle popolazioni di pipistrelli, e dei banchi di pesci, solo per citarne alcune.

Gli algoritmi SI possono essere di due tipologie:

- bio-inspired, come la Particle Swarm Optimization (PSO), il Fish Schooling Search (FSS), il firefly algorithm-I, il firefly algorithm-II, il buccock algorithm, il bat algorithm, l'artificial bee algorithm (ABC), il bacterial foraging optimization (BFO);
- non bio-inspired come il Fireworks algorithm (FWA), il water drops algorithm, brain storm optimization (BSO), e il magnetic optimization algorithm.

## 2.3 PSO-Particle Swarm Optimization

La Particle Swarm Optimization (PSO) è una metaeuristica di ispirazione biologica e population-based, introdotta nel 1995 da Kennedy e Eberhart. La PSO appartiene alla classe

---

<sup>10</sup> La Computazione Evolutiva, in inglese Evolutionary Computation, è uno dei sottocampi della Computational Intelligence, ossia l'intelligenza computazionale. La Computational Intelligence è riferita all'abilità di un computer nell'imparare compiti specifici da alcuni dati o da un'osservazione sperimentale: alcuni campi della Computational Intelligence sono le Reti Neurali Artificiali, la logica Fuzzy, la Chaos Computation, il Simulated Annealing, la Tabu Search, e, come già accennato nel paragrafo, la Computazione Evolutiva.

degli algoritmi Swarm Intelligence e s'ispira al comportamento degli stormi di uccelli nel cercare il mais. L'idea di Kennedy e Eberhant è stata quella di sfruttare le interazioni sociali piuttosto che le abilità individuali di ogni singolo. Nella PSO gli individui sono chiamati particelle, e, per analogia, come lo stormo di uccelli cerca il mais, così l'insieme delle particelle cerca l'ottimo della funzione obiettivo o funzione di fitness. Ogni particella è quindi situata in uno spazio, detto appunto spazio di ricerca di un problema o di una funzione; in questo caso si tratta di un problema di minimizzazione, ossia:

$$\min f(x) \quad \text{con } x \in R^d \quad (2.5)$$

dove  $f: R^d \rightarrow R$  è la funzione obiettivo del problema di minimizzazione.

Ogni particella è caratterizzata da tre vettori, ognuno di dimensione  $d$ , dove  $d$  sono le dimensioni dello spazio di ricerca. Se si considera una popolazione di  $M$  particelle in uno spazio multidimensionale, all'iterazione  $k$  – *esima*, allora i tre vettori della particella  $j$ -esima, dove  $j \in \{1, \dots, M\}$ , sono:

- $X_j^k \in R^d$ , è la posizione della  $j$ -esima particella alla  $k$ -esima iterazione, rappresentata in uno spazio  $d$ -dimensionale:

$$X_j^k = (x_{j,1}^k, x_{j,2}^k, \dots, x_{j,d}^k); \quad (2.6)$$

- $V_j^k \in R^d$  è la velocità della  $j$ -esima particella alla  $k$ -esima iterazione, rappresentata in uno spazio  $d$ -dimensionale:

$$V_j^k = (v_{j,1}^k, v_{j,2}^k, \dots, v_{j,d}^k); \quad (2.7)$$

- $P_j \in R^d$  è la migliore posizione in cui la particella è stata finora.

Le  $M$  particelle a ogni iterazione valutano la funzione obiettivo; la migliore posizione di ogni particella sarà quella associata al valore più basso della funzione di fitness. Se si pensa al problema di selezione del portafoglio, si può affermare che il vettore  $X_j^k$  non è altro che il vettore composto dai pesi dei vari titoli azionari, individuato dalla  $j$ -esima particella alla  $k$ -esima iterazione; di conseguenza la miglior posizione di ogni singola particella individuata finora,  $P_j$ , non è altro che il vettore dei pesi dei titoli, associato al più basso valore della funzione fitness, valutata dalla  $j$ -esima particella.

## 2.4 Le modifiche alla PSO

Negli anni, la PSO ha subito molti cambiamenti dalla sua introduzione, soprattutto per quanto riguarda l'equazione della velocità  $V_j^k$ , compresi i relativi metodi di controllo di essa, e la topologia considerata nella connessione sociale. Di seguito si proporranno la versione originale di Kennedy e Eberhart della PSO, e alcune successive modifiche.

### Algoritmo originale

L'algoritmo originale della PSO prevede i seguenti passi:

- 1) Inizializzazione della popolazione, tramite una matrice di particelle con posizioni e velocità casuali nello spazio di ricerca di dimensioni  $d$ .
- 2) ciclo che continua fino a quando il criterio di arresto scelto non è soddisfatto:
  - a. impostare  $k = 1$  e valutare la funzione obiettivo  $f(X_j^k)$  per  $j = 1, \dots, M$ ;
  - b. se  $f(X_j^k) \leq P_j$  allora impostare  $P_j = X_j^k$ . In questo punto il valore della funzione di fitness alla  $k$ -esima iterazione viene valutato con il valore della funzione finora migliore. Se il valore corrente della funzione è migliore, allora la posizione della particella associata a tale valore viene settata come la nuova migliore posizione della  $j$ -esima particella;
  - c. identificare la particella nel vicinato con la migliore posizione e assegnarle l'indice  $g$ , ossia  $P_{g(j)}$ .
  - d. Aggiornare la posizione e la velocità della particella secondo la seguente equazione:

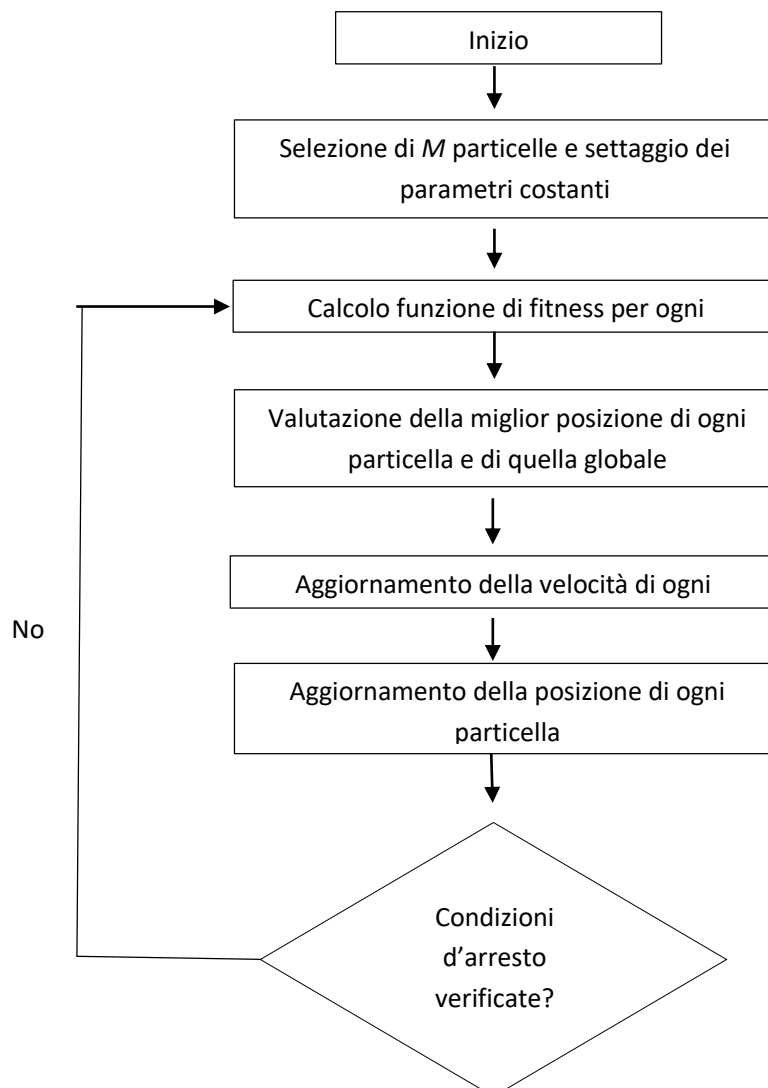
$$\begin{cases} V_j^{k+1} = V_j^k + U(0, \phi_1) \otimes (P_j - X_j^k) + U(0, \phi_2) \otimes (P_{g(j)} - X_j^k) & (2.8) \\ X_j^{k+1} = X_j^k + V_j^k & (2.9) \end{cases}$$

- e. Se si soddisfa uno dei criteri, come ad esempio una funzione di fitness sufficientemente buona o un numero massimo di iterazioni, uscire dal ciclo continuo, altrimenti rintonare al punto  $b$ .

L'equazione (2.8) si può vedere come la velocità dell'iterazione successiva,  $V_j^{k+1}$ , è calcolata come somma di tre componenti:

- $V_j^k$  è la velocità corrente della  $j$ -esima particella;
- $U(0, \phi_1) \otimes (P_j - X_j^k)$  è la differenza tra la migliore posizione della  $j$ -esima particella e la posizione corrente. Il tutto è moltiplicato per un vettore di numeri casuali estratti da una distribuzione uniforme  $[0, \phi_1]$  a ogni iterazione e per ogni particella. Quindi la particella è attratta nella direzione della sua miglior posizione;
- $U(0, \phi_2) \otimes (P_{g(j)} - X_j^k)$  è la differenza tra la posizione migliore del vicinato<sup>11</sup>, trovata finora, e la posizione corrente della particella, moltiplicata per un vettore di numeri casuali estratti da una distribuzione uniforme  $[0, \phi_2]$  a ogni iterazione e per ogni particella. Questa componente della velocità fa sì che la particella sia attratta verso la migliore posizione del vicinato trovata finora.

L'algoritmo appena descritto può essere rappresentato in un diagramma di flusso come segue:



<sup>11</sup> il vicinato può variare a seconda della topologia scelta. Questo argomento verrà trattato nel paragrafo 2.5.

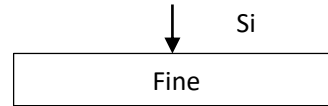


Figura 2.1: diagramma di flusso della PSO

Per quanto riguarda il controllo della velocità e il settaggio dei parametri nella prima versione dell'algoritmo si è proposto innanzitutto di individuare la dimensione della popolazione in un range compreso tra 20 e 50 particelle.

I parametri  $\phi_1$  e  $\phi_2$ , detti coefficienti di accelerazione, rappresentano la 'forza casuale' nella direzione rispettivamente della miglior posizione personale della particella e della miglior posizione del vicinato. Il settaggio sbagliato di questi due parametri può portare ad una instabilità della stessa PSO e a una velocità che aumenta senza alcun controllo. Nei primi studi fu utilizzato il valore  $\phi_1 = \phi_2 = 2$  che rendeva instabile la PSO. Per controllare la velocità fu quindi posto un range  $[-V_{max}, +V_{max}]$  per ogni velocità  $V_j^k$ . Questa soluzione sebbene da un lato controlla la velocità, dall'altro la stessa scelta di  $V_{max}$  è complessa in quanto questo parametro influenza l'equilibrio tra sfruttamento e esplorazione. Inoltre non esiste alcuna regola da seguire per settare in modo corretto la velocità massima. Per ovviare al problema sono state proposte altre soluzioni per controllare la velocità.

### Modifica 1: peso d'inerzia

Questa prima modifica fu proposta da Shi e Eberhart nel 1998:

$$\begin{cases} V_j^{k+1} = \omega V_j^k + U(0, \phi_1) \otimes (P_j - X_j^k) + U(0, \phi_2) \otimes (P_{g(j)} - X_j^k) \\ X_j^{k+1} = X_j^k + V_j^k \end{cases} \quad (2.10)$$

$$(2.11)$$

dove  $\omega$  è il peso d'inerzia.

Mentre le due ultime componenti della velocità rappresentano le forze di attrazione esterne, quindi lo sfruttamento verso la miglior posizione personale della particella e verso la miglior posizione del vicinato, la componente  $\omega V_j^k$  rappresenta invece l'esplorazione della particella.

Più alto è il valore assegnato al peso d'inerzia e più la velocità dell'iterazione successiva dipende dalla velocità attuale  $V_j^k$  della particella. Viceversa se il valore assegnato al peso d'inerzia è basso, allora la velocità  $V_j^{k+1}$  sarà influenzata più dalle due componenti di sfruttamento che dalla velocità corrente della particella. Una delle possibili soluzioni per quanto riguarda il valore del peso d'inerzia potrebbe essere settare inizialmente un valore elevato di  $\omega$  in modo da privilegiare l'esplorazione allo sfruttamento, e in seguito assegnare valori decrescenti al peso d'inerzia, preferendo lo sfruttamento, quindi l'attrazione verso l'ottimo locale, rispetto all'esplorazione. Una delle formule proposte prevede appunto che il parametro  $\omega$  diminuisca con l'aumentare delle iterazioni.

$$\omega^k = \omega_{max} + \frac{\omega_{min} - \omega_{max}}{K} k, \quad (2.12)$$

dove

$\omega_{max}$  e  $\omega_{min}$  sono comunemente settati rispettivamente a 0,9 e 0,4;

$K$  è il numero massimo di iterazioni consentite;

$k$  è il numero dell'iterazione corrente.

Un'altra soluzione proposta prevede l'uso del peso d'inerzia con una componente casuale, che decresce con il tempo. Ad esempio Eberhart e Shi nel 2001 hanno utilizzato  $\omega = U(0.5,1)$ .

Il corretto settaggio del peso d'inerzia e dei coefficienti di accelerazione rendono quindi la PSO più stabile rispetto al primo modello, al punto da poter assegnare un valore alto al parametro  $V_{max}$ , o addirittura non settarlo.

### **Modifica 2: limitazione dei coefficienti**

Il metodo della limitazione dei coefficienti fu introdotto da Clerc e Kennedy nel 2002. La limitazione dei coefficienti assicura la convergenza, e previene da un incremento non controllato della velocità. La nuova equazione per la velocità è così composta,

$$\begin{cases} V_j^{k+1} = \chi \left( V_j^k + U(0, \phi_1) \otimes (P_j - X_j^k) + U(0, \phi_2) \otimes (P_{g(j)} - X_j^k) \right) \\ X_j^{k+1} = X_j^k + V_j^k \end{cases} \quad (2.13)$$

$$(2.14)$$

dove,

$$\chi = \frac{2}{\phi - 2 + \sqrt{\phi^2 - 4\phi}} \quad \text{con } \phi = \phi_1 + \phi_2 > 4 \quad (2.15)$$

In Clerk e Kennedy il valore di  $\phi$  è posto pari a 4.1, con  $\phi_1 = \phi_2$ ; sostituendo nella (2.16) si ottiene che il moltiplicatore costante  $\chi$  è approssimativamente 0.7298. Quindi, tornando all'equazione (2.13), si può affermare che la velocità corrente è moltiplicata per 0.7298, mentre le due componenti associate allo sfruttamento sono moltiplicate per un numero casuale compreso nell'intervallo  $[0, 1.496118]$ ; infatti poiché  $\phi_1 = \phi_2 = 2.05$ , allora  $0.7298 \times 2.05 \sim 1.49618$ .

Si può notare come la PSO con il peso d'inerzia può essere equivalente alla PSO con la limitazione dei coefficienti se si pone  $\omega = \chi$  e  $\phi_i = \chi\phi_i$ . In questo modo i valori nella PSO con il peso d'inerzia saranno  $\omega = 0.7298$  e  $\phi_1 = \phi_2 = 1.49618$ .

### Modifica 3: Fully Informed Particle Swarm (FIPS)

Kennedy e Mendes nel 2002 hanno rivisto come la particella interagisce con il vicinato. Mentre nella versione originale della PSO la particella mantiene solo due informazioni, ossia la sua miglior posizione e la migliore posizione del vicinato, nella versione di Kennedy e Mendes la particella mantiene tutta l'informazione, ossia le posizioni di tutti i suoi vicini. Il metodo FIPS è così sviluppato:

$$\begin{cases} V_j^{k+1} = \chi \left( V_j^k + \frac{1}{K_j} \sum_{n=1}^{K_j} U(0, \phi) \otimes (P_{nbr_n^j} - X_j^k) \right) \\ X_j^{k+1} = X_j^k + V_j^k \end{cases} \quad (2.17)$$

$$(2.18)$$

dove

$K_j$  è il numero dei vicini della particella  $j$ -esima;

$nbr_n^j$  è l' $n$ -esimo vicino della particella  $j$ -esima.

Si nota che se il numero dei vicini della particella  $j$ -esima viene posto uguale a 2,  $K_j = 2$ , allora il modello risulta uguale alla PSO tradizionale. Dai risultati di Kennedy e Mendes si evince che questa versione della PSO trova una migliore soluzione in meno iterazioni rispetto all'algoritmo tradizionale, tuttavia dipende in maniera pesante dalla topologia della popolazione.

## 2.5 Topologia della popolazione

La topologia definisce la struttura del vicinato e il tipo di connessione che esiste tra le particelle della popolazione. Le topologie possono essere divise in due macro-categorie:

- topologie statiche, dove i vicini e il vicinato non cambiano con l'aumentare dell'iterazioni. Le principali topologie statiche sono:
  - la topologia *lbest*, ossia *local best*. Kennedy e Eberhart hanno utilizzato la topologia *lbest* nell'algoritmo originale della PSO, impostando  $K = 2$ , ossia ogni particella è connessa con 2 particelle adiacenti in una struttura ad anello; infatti nello specifico caso in cui  $K = 2$  si può parlare di *ring topology*, ossia topologia ad anello, rappresentata nella Figura 2.2. Sebbene la topologia *lbest* converga lentamente alla soluzione ottima è meno vulnerabile all'attrazione al minimo locale rispetto ad altre topologie, come ad esempio la topologia *gbest*.

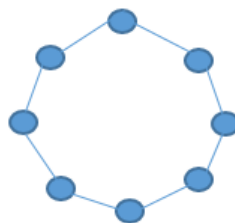


Figura 2.2: rappresentazione ring topology

- la topologia *gbest*, ossia *global best*. In questo caso l'intera popolazione è usata come vicinato di ogni particella, come si può vedere nella Figura 2.3.

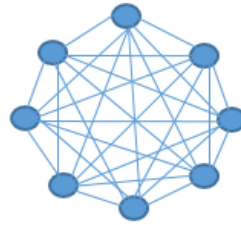


Figura 2.3: rappresentazione *gbest* topology

- la topologia di von Neumann, dove ogni particella è collegata con la particella sopra, con quella sotto, con quella alla propria sinistra e con quella alla propria destra. Kennedy e Mendes nel 2002 e Mendes nel 2004 hanno investigato l'impatto di diverse topologie nella performance dell'algoritmo. La topologia di von Neumann è risultata essere la migliore, in quanto con  $K=4$ , la particella è da una parte connessa con un numero maggiore di particelle rispetto alla topologia *lbest*, ma dall'altra la densità è minore rispetto alla topologia *gbest*. Per quanto invece riguarda la topologia *gbest* applicata all'algoritmo tradizionale della PSO, si è osservato che è stata la più veloce a convergere ma con i peggiori risultati. Inoltre si è visto che la PSO FIPS dà risultati superiori rispetto alle altre versioni della PSO.

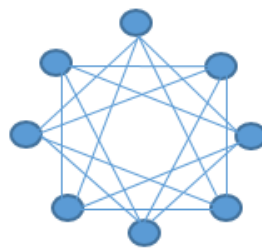


Figura 2.4: rappresentazione von Neumann topology

- topologie dinamiche, dove invece i vicini e i vicinati cambiano con l'aumentare delle iterazioni. In particolare:
  - nel 1999 Suganthan propose d'iniziare la ricerca con una topologia *lbest* con struttura ad anello e incrementare lentamente con l'aumentare del numero di iterazioni la dimensione del vicinato fino ad avere una topologia *gbest* alla fine

delle iterazioni. Il motivo di questa scelta sta nel fatto di poter sfruttare inizialmente la miglior esplorazione della topologia *lbest*, e solo successivamente sfruttare invece la veloce convergenza della topologia *gbest*.

- nel 2003 Peram et al. hanno utilizzato una distanza Euclidea pesata per identificare il vicino con cui la particella deve interagire. Lo scopo è quello di limitare la convergenza prematura della PSO tradizionale. In sostanza l'algoritmo proposto da Peram et al. spinge le particelle vicino a quelle con il migliore valore della funzione di fitness. La velocità di ogni particella viene aggiornata selezionando una particella che massimizza il rapporto tra la differenza delle fitness e la distanza per ogni dimensione. In altre parole la *d-esima* dimensione della particella *j-esima* è aggiornata usando una particella chiamata *nbest*, con la migliore posizione  $P_i$  scelta per minimizzare il seguente rapporto:

$$\frac{Fitness(P_i) - Fitness(X_j)}{|P_{id} - X_{jd}|} \quad (2.19)$$

L'espressione (2.19) è chiamata Fitness-Distance-Ratio.

- nel 2005 Jason e Middenford hanno utilizzato la topologia *hierarchy*. In questa topologia ogni particella è influenzata dalla sua miglior posizione passata e dalla particella direttamente sopra di essa in una lista in cui le particelle sono elencate secondo le loro prestazioni. Le particelle con le miglior prestazioni si muovono verso l'alto nella gerarchia;
- nel 2006 Clerc ha introdotto la topologia TRIBES, dove le caratteristiche della topologia, come la dimensione della popolazione, cambiano durante le iterazioni a seconda dei risultati. In pratica la popolazione è divisa in sottopopolazioni: nelle popolazioni migliori le particelle peggiori vengono scartate; quest'ultime vengono aggiunte alle popolazioni peggiori, aumentando così la possibilità di quest'ultime di migliorare.

## 2.6 FWA-Fireworks Algorithm

Il Fireworks Algorithm (FWA) è un algoritmo appartenente alla classe di algoritmi Swarm Intelligence. Introdotto nel 2010 da Y. Tan e Y. Zhu, il FWA trae ispirazione dal processo di

esplosione dei fuochi d'artificio. Quando un fuoco d'artificio esplode, lo spazio circostante viene riempito dalle scintille (sparks): l'esplosione può essere vista come la ricerca nello spazio locale attorno al fuoco d'artificio.

Il FWA è costituito da quattro componenti chiave, detti building blocks, che verranno analizzati in questo capitolo: explosion operator, mutation operator, mapping rule e selection strategy.

Prima di analizzare in dettaglio le fasi dell'algoritmo, si ricorda che il problema di minimizzazione è analogo a quello tipicamente affrontato nella PSO:

$$\min f(x) \in R, x_{min} \leq x \leq x_{max} \quad 2.20$$

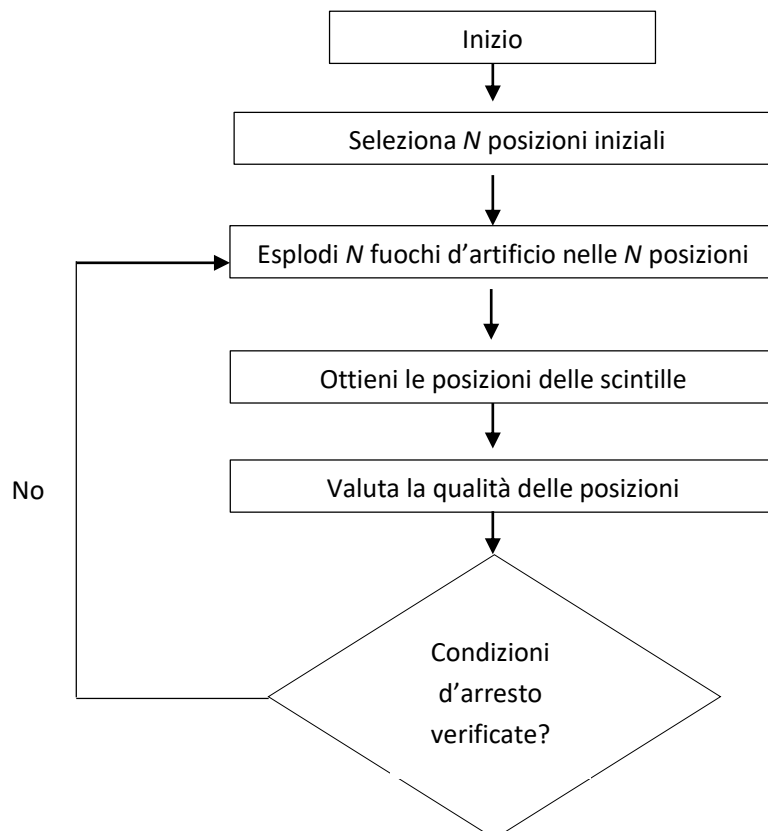
dove  $x = x_1, x_2, \dots, x_d$  è la posizione del fuoco d'artificio nello spazio di ricerca,

$f(x)$  è la funzione obiettivo o funzione di fitness,

$x_{min}$  e  $x_{max}$  sono rispettivamente il limite inferiore e superiore dello spazio di ricerca.

## 2.7 Il funzionamento del FWA

Il processo del Fireworks Algorithm può essere rappresentato nel seguente diagramma di flusso.



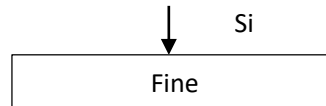


Figura 2.5: diagramma di flusso del FWA

Il primo passo dell'algoritmo, come si può vedere dalla Figura 2.5, è quello di selezionare  $N$  posizioni iniziali e generare  $N$  fuochi d'artificio casualmente. Questi fuochi d'artificio sono fatti "esplodere" nelle  $N$  posizioni selezionate. È in questa fase che entra in gioco l'explosion operator che include l'explosion strenght, l'explosion amplitude e il displacement operation, che verranno descritti nel paragrafo successivo.

Si specifica che le posizioni dei fuochi d'artificio costituiscono l'insieme della popolazione delle possibili soluzioni. Ogni fuoco d'artificio quando "esplode" produce un determinato numero di scintille, ognuna con la sua posizione. Dopo l'esplosione si avranno sia le posizioni dei fuochi d'artificio iniziale e sia quelle delle scintille prodotte. Da questo insieme di posizioni è selezionato, attraverso due criteri che saranno analizzati in seguito, un numero di posizioni uguale alla popolazione iniziale di fuochi d'artificio. Questa popolazione con nuove posizioni può costituire la soluzione al problema se soddisfa i criteri d'arresto, altrimenti sarà la nuova popolazione iniziale da cui far ripartire l'algoritmo.

### 2.7.1 Explosion Strenght

Nell'explosion strenght il numero di scintille generato da ogni fuoco d'artificio è determinato come segue:

$$S_i = m * \frac{Y_{max} - f(x_i) + \varepsilon}{\sum_{i=1}^n (Y_{max} - f(x_i)) + \varepsilon} \quad (2.21)$$

dove

$S_i$  è il numero di scintille di ogni fuoco d'artificio;

$m$  è una costante che controlla il numero totale di scintille;

$f(x_i)$  è il valore della funzione obiettivo che assume il fuoco d'artificio  $x_i$ ;

$Y_{max} = \max(f(x_i))$  con  $i = 1, \dots, N$  è il massimo valore che assume la funzione obiettivo tra gli  $N$  fuochi d'artificio (poiché il problema è di minimizzazione della funzione obiettivo,  $Y_{max}$  sarà il peggior valore che  $f(x_i)$  può assumere);

$\varepsilon$  è un numero costante di grandezza trascurabile, utilizzato per evitare che il numeratore e il denominatore assumano il valore zero.

Per evitare che le scintille prodotte dai fuochi d'artificio siano generate in numero troppo elevato, sono definiti dei vincoli per  $S_i$ ,

$$\hat{s}_i = \begin{cases} \text{round}(a \cdot m) & \text{se } S_i < am \\ \text{round}(b \cdot m) & \text{se } S_i > bm, \text{ con } a < b < 1 \\ \text{round}(s_i) & \text{altrimenti} \end{cases} \quad (2.22)$$

dove

$\hat{s}_i$  è la limitazione al numero di scintille;

$a$  e  $b$  sono costanti;

$\text{round}()$  è una funzione che permette di arrotondare al numero di cifre specificato, in questo caso all'intero.

Analizzando l'equazione (2.21) si nota che i fuochi d'artificio con una miglior funzione di fitness generano un numero maggiore di scintille e viceversa. Se il denominatore rimane costante, il numeratore sarà maggiore per fuochi d'artificio con una funzione di fitness minore, determinando così una generazione maggiore di scintille rispetto agli altri fuochi d'artificio con un valore della funzione di fitness più elevato.

### 2.7.2 Explosion amplitude

Dopo aver calcolato il numero di scintille, l'algoritmo calcola l'ampiezza di ogni esplosione. L'explosion amplitude è definita come:

$$A_i = \hat{A} * \frac{f(x_i) - Y_{min} + \varepsilon}{\sum_{i=1}^n (f(x_i) - Y_{min}) + \varepsilon} \quad (2.23)$$

dove

$A_i$  è l'ampiezza di ogni fuoco d'artificio;

$\hat{A}$  è una costante che indica la somma di tutte le ampiezze;

$Y_{min} = \min(f(x_i))$  con  $i = 1, \dots, N$  è il valore minimo che assume la funzione obiettivo, quindi il valore migliore;

$f(x_i)$  è il valore della funzione obiettivo per il fuoco d'artificio  $x_i$ ;

$\varepsilon$  è un numero costante di grandezza trascurabile, utilizzato per evitare che il numeratore e il denominatore assumano il valore zero.

Come si può vedere dalla formula (2.23) i fuochi d'artificio con una migliore funzione di fitness avranno un'ampiezza minore e viceversa. Infatti se il denominatore rimane costante, il numeratore diminuisce al diminuire del valore della funzione di fitness, e quindi di conseguenza l'ampiezza sarà minore per fuochi d'artificio con una funzione di fitness migliore.

Si può quindi concludere che la posizione del fuoco d'artificio determina il numero delle scintille e l'ampiezza, identificando così due tipologie di fuochi d'artificio:

- fuochi d'artificio di buona qualità: questi fuochi d'artificio generano una grande popolazione di scintille in un'ampiezza ridotta;
- fuochi d'artificio di cattiva qualità: questi fuochi d'artificio generano invece una piccola popolazione di scintille sparpagliate entro una grande ampiezza.

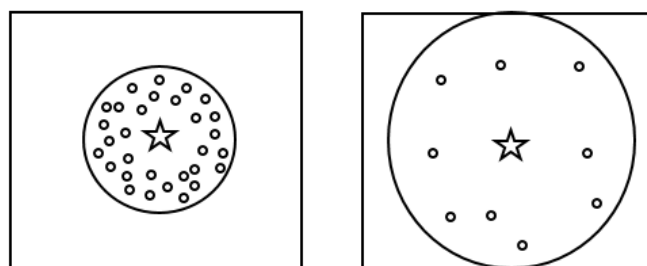


Figura 2.6: a sinistra è rappresentata l'esplosione di un fuoco d'artificio di buona qualità, mentre a destra uno di scarsa qualità.

Le due tipologie di fuoco d'artificio servono per due scopi differenti. I fuochi d'artificio con un alto valore della funzione di fitness (cosa non buona poiché il problema di ottimizzazione è un

problema di minimizzazione), esplorano lo spazio di ricerca, evitando che l'algoritmo converga troppo velocemente all'ottimo, mentre i fuochi d'artificio migliori, producendo molte scintille in una piccola zona, adempiono alla funzione di sfruttamento, ossia ricercano in modo approfondito in una zona promettente dello spazio di ricerca: quindi più la posizione del fuoco d'artificio è vicina alla soluzione e più l'algoritmo genera scintille in un'ampiezza ridotta.

### 2.7.3 Displacement operation

L'ultima fase dell'esplosione si conclude con il displacement operation, ossia "un'operazione di spostamento". È proprio lo spostamento, indicato con l'operatore  $h$ , che permette di generare le posizioni del numero di scintille determinato nella fase dell'explosion strenght. Tuttavia questo spostamento non è applicato a tutte le dimensioni del fuoco d'artificio. Se si considera infatti che la posizione del fuoco d'artificio è  $d$ -dimensionale, la displacement operation viene applicata solo a  $z < d$  dimensioni scelte casualmente tra le  $d$  dimensioni. Le  $z$  dimensioni sono scelte come segue,

$$z = \text{round}(d \cdot \text{rand}(0,1)) \quad (2.24)$$

dove

$d$  sono le dimensioni della posizione  $x_i$  del fuoco d'artificio

$\text{rand}(0,1)$  è una distribuzione uniforme nell'intervallo  $[0,1]$ ,

$\text{round}()$  è una funzione che permette di arrotondare al numero di cifre specificato.

Lo spostamento  $h$  per ogni dimensione  $z$  è calcolato come segue,

$$h = A_i \cdot \text{rand}(-1,1) \quad (2.25)$$

dove

$A_i$  è l'ampiezza dell'esplosione associato alla posizione  $x_i$ ,

$\text{rand}(-1,1)$  è una distribuzione uniforme nell'intervallo  $[0,1]$ .

Quindi la nuova posizione della  $j$ -esima dimensione tra le  $z$  selezionate sarà calcolata come:

$$x_i^{j+1} = x_i^j + h \quad (2.26)$$

dove

$x_i^{j+1}$  è la nuova posizione della  $j$ -esima dimensione,

$x_i^j$  è l'attuale posizione della  $j$ -esima dimensione,

#### 2.7.4 Mutation Operator

Il secondo building block dell'algorithm FWA è il mutation operator, che garantisce la diversità della popolazione attraverso ulteriori esplosioni che generano nuove scintille dagli  $N$  fuochi d'artificio. Il mutation operator utilizza il Gaussian operator per generare le nuove scintille; queste sono posizionate tra il miglior fuoco d'artificio e i fuochi d'artificio selezionati. In questo caso l'ampiezza e il numero di scintille generate non sono influenzate dalla posizione del fuoco d'artificio. Il numero delle scintille è infatti posto pari alla costante  $\hat{m}$ , mentre l'ampiezza è casuale. Anche in questo caso sono selezionate  $z$  dimensioni attraverso la l'equazione (2.24). La nuova posizione della  $j$ -esima dimensione sarà calcolata come:

$$x_i^{j+1} = x_i^j \cdot g \quad (2.27)$$

dove

$x_i^{j+1}$  è la nuova posizione della  $j$ -esima dimensione,

$x_i^j$  è l'attuale posizione della  $j$ -esima dimensione,

$g$  è un numero casuale estratto dalla distribuzione Gaussiana con media 1 e varianza 1:

$$g = \text{Gaussian}(1,1). \quad (2.28)$$

L'equazione (2.27) è ripetuta per tutte le  $z$  dimensioni e per il numero di scintille  $\hat{m}$ .

### 2.7.5 Mapping Rule

Se la posizione di un fuoco d'artificio è vicino al limite dello spazio di ricerca, l'esplosione potrebbe generare scintille che cadono in parte fuori dallo spazio ammesso. Poiché le scintille fuori dalla regione ammissibile sono inutili la mapping rule interviene mappando nuovamente le scintille dentro lo spazio di ricerca.

La mapping rule utilizza l'aritmetica modulare e viene indicata come,

$$x_i^k = X_{LB,k} + x_i^k \% (X_{UB,k} - X_{LB,k}) \quad (2.29)$$

dove,

$x_i^k$  rappresenta la posizione di ogni scintilla che giace fuori dai confini;

$X_{LB,k}$  è il low boundary, ossia il limite inferiore della posizione di una scintilla;

$X_{UB,k}$  è il upper boundary, ossia il limite superiore della posizione di una scintilla.

$\%$  è il simbolo dell'aritmetica modulare.

Per comprendere meglio come funziona la mapping rule si propone il seguente esempio: lo spazio di ricerca è compreso tra -120 e +120 e l'ampiezza massima è impostata a 50. Se due fuochi d'artificio sono posizionati ai due limiti, le scintille potrebbero cadere negli intervalli [-170, -120] e [+120,+170]. A questo punto la mapping rule mappa le due posizioni nell'intervallo [0,50] per evitare che le scintille ricadano fuori dalla regione ammissibile.

La Figura (2.7) esemplifica il caso appena illustrato.

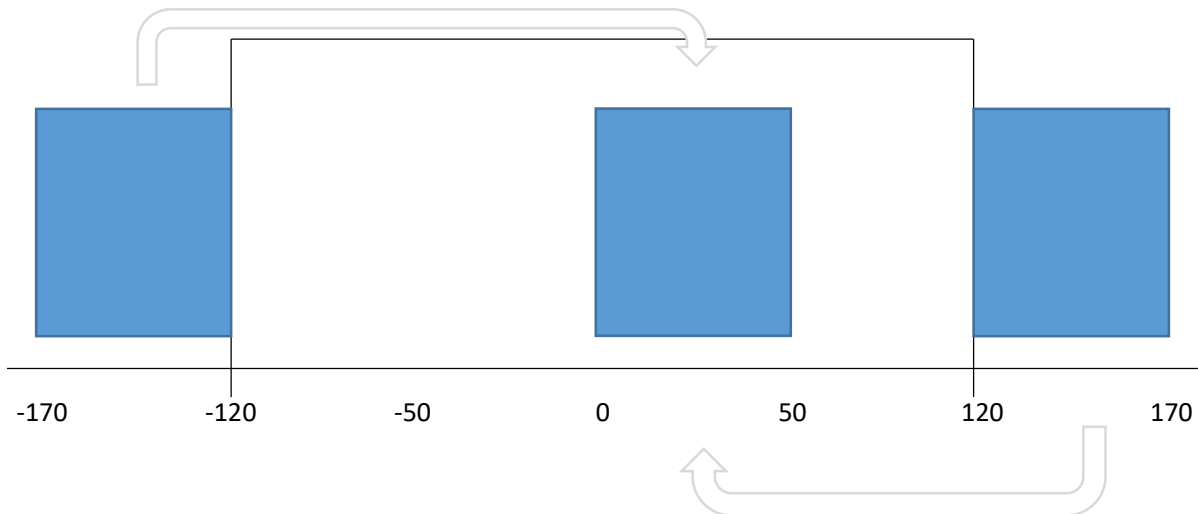


Figura 2.7: Funzionamento della mapping rule

### 2.7.6 Selection strategy

La selection strategy serve per selezionare le scintille che saranno utilizzate per individuare le nuove posizioni dei fuochi d'artificio nella successiva iterazione. Questa fase avviene se la generazione corrente non soddisfa i criteri d'arresto, che possono essere ad esempio un numero massimo di iterazioni o ad esempio, nel caso specifico di selezione di portafogli di titoli azionari, il raggiungimento di un rendimento atteso prefissato. Quindi se i criteri di arresto non sono soddisfatti, l'algoritmo seleziona le nuove posizioni utilizzando una strategia di selezione chiamata distance-based strategy. Al fine di selezionare le scintille per la prossima generazione, la scintilla con la migliore posizione,  $x^*$ , è sempre utilizzata nell'iterazione successiva. La miglior posizione tra tutte quelle individuate è quella con la funzione di fitness,  $f(x^*)$ , più bassa. Poiché il numero di fuochi d'artificio rimane costante per ogni iterazione, rimangono  $(N - 1)$  posizioni da selezionare attraverso la distance-based strategy. Questa strategia consiste nel selezionare le scintille in base alla distanza tra ogni posizione e le altre. La distanza tra una posizione  $x_i$  e le altre posizioni è calcolata come segue:

$$R(x_i) = \sum_{j=1}^N d(x_i, x_j) = \sum_{j=1}^N \|x_i, x_j\| \quad (2.30)$$

dove

$R(x_i)$  è la somma delle distanze tra  $x_i$  e le altre scintille,

$d(x_i, x_j)$  è la distanza Euclidea tra le due posizioni  $x_i$  e  $x_j$ ,

l'indice  $j \in K$  significa che la posizione  $j$  appartiene a  $K$ , con  $K$  insieme di tutte le scintille, sia quelle generate dall'explosion operator e sia quelle generate dal mutation operator. La distanza così calcolata viene utilizzata per calcolare una “probabilità” di scegliere la scintilla  $x_i$  è definita come,

$$p(x_i) = \frac{R(x_i)}{\sum_{j \in K} R(x_j)} \quad (2.31)$$

Dall'equazione (2.31) si può vedere che le scintille con una maggior distanza hanno più probabilità di essere scelte per la generazione successiva, in modo da garantire la diversità della popolazione.

### 2.7.7 Pseudo-codice FWA

In conclusione si riporta uno pseudo-codice del FWA

Pseudo-codice FWA

- 
- 1: Inizializza  $N$  posizioni e calcola la funzione di fitness  $f(x_i)$  per ogni fuoco d'artificio;
  - 2: **fino a che** le condizioni di arresto non sono raggiunte
  - 3: **per** ogni posizione  $x_i = 1:N$
  - 4:     Calcola la funzione di fitness  $f(x_i)$ ;

- 5: Calcola il numero di scintille  $S_i$ ;
- 6: Calcola l'ampiezza dell'esplosione  $A_i$ ;
- 8: Seleziona casualmente  $z$ -dimensioni della posizione  $x_i$ ;  
%displacement operation
- 9: **per** ogni dimensione  $x_i^j \in z$  dimensioni di  $x_i$
- 10: Calcola lo spostamento:  $h = A_i \cdot rand(-1,1)$ ;
- 11: Calcola la nuova posizione  $x_i^{j+1} = x_i^j + h$  ;
- 12: **se**  $x_i^{j+1} < x_i^{min}$  o  $x_i^{j+1} > x_i^{max}$  allora
- 13: mappa  $x_i^{j+1}$  nello spazio potenziale:  $x_i^k = X_{LB,k} + x_i^k \% (X_{UB,k} - X_{LB,k})$
- 14: **fine ciclo se**
- 15: **fine ciclo per**  
%mutation operator
- 16: **per**  $k = 1: \hat{m}$
- 16: Seleziona casualmente  $z$ -dimensioni della posizione  $x_i$ ;
- 17: **per** ogni dimensione  $x_i^j \in z$  dimensioni di  $x_i$
- 18: Calcola il coefficiente dell'esplosione Gaussiana:  $g = Gaussian(1,1)$ ;
- 17: Calcola le scintille dell'esplosione Gaussiana:  $x_i^{j+1} = x_i^j \cdot g$ ;
- 19: **se**  $x_i^{j+1} < x_i^{min}$  o  $x_i^{j+1} > x_i^{max}$  allora
- 20: mappa  $x_i^{j+1}$  nello spazio potenziale:  $x_i^k = X_{LB,k} + x_i^k \% (X_{UB,k} - X_{LB,k})$
- 21: **fine ciclo se**

22: **fine del ciclo per**

23: **fine ciclo per**

24: Seleziona la posizione della scintilla con il più basso valore della funzione di fitness e le

altre posizioni secondo la probabilità:  $p(x_i) = \frac{R(x_i)}{\sum_{j \in K} R(x_j)}$

25: **fine del ciclo fino a che**

## 2.8 Caratteristiche del Fireworks Algorithm

Alla luce del funzionamento del FWA appena descritto, si possono individuare alcune caratteristiche che distinguono il FWA.

- **Esplosività:** dopo la prima iterazione, i fuochi d'artificio esplodono generando un numero di scintille. La selection strategy individua  $N$  scintille che saranno le nuove posizioni dei fuochi d'artificio dell'iterazione successiva.
- **Istantaneità:** le scintille hanno "vita breve", ossia quelle che non vengono selezionate svaniscono.
- **Semplicità:** ogni fuoco d'artificio ha solo la sua informazione.
- **Località:** ogni fuoco d'artificio genera scintille all'interno della propria ampiezza.
- **Proprietà di emergenza della complessità:** è una proprietà che un sistema complesso ha, ma che gli individui che lo formano non hanno. I fuochi d'artificio infatti collaborano tra di loro e come gruppo mostrano un comportamento più complesso di quello che mostrano i singoli.
- **Parallelismo:** non c'è un meccanismo centrale che controlla i fuochi d'artificio; i fuochi d'artificio sono infatti indipendenti.
- **Diversità:** questa caratteristica è fondamentale per l'algoritmo. Infatti grazie alla diversità, l'algoritmo può convergere all'ottimo globale, senza rimanere 'intrappolato' nell'ottimo locale. La diversità si può individuare nel numero di scintille e nell'ampiezza delle esplosioni che cambiano a seconda della posizione del fuoco d'artificio. Inoltre l'algoritmo ha due operatori che garantiscono la diversità, il displacement operator e il gaussian mutation. Anche nella strategia di selezione si può riscontrare il mantenimento della diversità: infatti le scintille più distanti dalle altre

hanno una probabilità maggiore di essere scelte, proprio per preservare la diversità della popolazione.

- Robustezza dell'algoritmo: poiché i fuochi d'artificio sono indipendenti tra di loro, i peggiori non influenzano pesantemente il comportamento dell'algoritmo.

## 2.9 Confronto tra PSO e FWA

In questo paragrafo si effettueranno alcuni confronti tra i due algoritmi analizzati, la PSO nella versione con il peso d'inerzia, e il FWA.

I due algoritmi fanno entrambi parte della classe di Swarm Intelligence, tuttavia la PSO è una metaeuristica bio-ispirata, mentre il FWA è una meta-euristica non bio-ispirata. La PSO e il FWA presentano un algoritmo in alcuni tratti simili e in altri diversi, come di seguito riportato:

1. La prima fase del Fireworks Algorithm inizializza  $N$  posizioni dei fuochi d'artificio, mentre la Particle Swarm Optimization inizializza  $M$  particelle.
2. Nel passo successivo tutti due gli algoritmi calcolano la funzione di fitness per ogni posizione: si parlerà di posizione del fuoco d'artificio per il FWA, e di posizione della particella per la PSO.
3. Nella terza fase la PSO confronta il valore della funzione di fitness associata alla posizione della particella della iterazione corrente con la miglior posizione passata: se la funzione di fitness corrente è minore, questa verrà considerata la nuova migliore posizione della particella; inoltre viene valutata la migliore posizione del vicinato. Infine vengono aggiornate la velocità e la posizione. Per quanto riguarda invece il FWA, il displacement operation e il mutation operation determinano le posizioni delle scintille. La miglior funzione di fitness associata alla miglior posizione viene memorizzata per l'iterazione successiva. La probabilità che le altre scintille hanno di essere selezionate per l'iterazione successiva dipende dalle reciproche distanze.

Le due strategie di selezione sono diverse anche per le teorie a cui si ispirano. Infatti il FWA s'ispira all'*immune concentration theory*. Ogni scintilla dell'algoritmo è 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 invece hanno una maggior distanza dalle altre. Questo meccanismo presente nel FWA preserva la diversità delle scintille, ossia della popolazione, meccanismo invece non presente nella PSO.

Per quanto riguarda invece i parametri, entrambi gli algoritmi hanno dei parametri iniziali da settare che rimangono costanti per tutte le iterazioni:

- Nella PSO le costanti da determinare sono:
  - i due coefficienti di accelerazioni,  $\phi_1$  e  $\phi_2$ ,
  - il numero di particelle  $M$ ,
  - $\omega_{max}$  e  $\omega_{min}$  per il calcolo del peso d'inerzia che decresce con il numero delle iterazioni,
  - $K$ , ossia il numero di iterazioni massimo;
  - $V_{max}$ , ossia la velocità massima.
- In FWA le costanti da impostare inizialmente sono:
  - il numero di fuochi d'artificio  $N$
  - $\hat{A}$ , ossia la massima ampiezza possibile per ogni fuoco d'artificio;
  - $x_{min}$  e  $x_{max}$ , limiti della regione di ricerca;
  - le dimensioni  $d$  di ogni posizione;
  - $K$ , ossia il numero di iterazioni massimo;
  - $\hat{m}$ , numero di scintille prodotte nell'esplosione Gaussiana.

I due parametri da settare comuni a tutti e due gli algoritmi sono la dimensione della popolazione e il massimo numero d'iterazioni.

Per cogliere maggiormente la differenza tra PSO e FWA si riportano due grafici che esemplificano il movimento delle particelle e delle scintille.

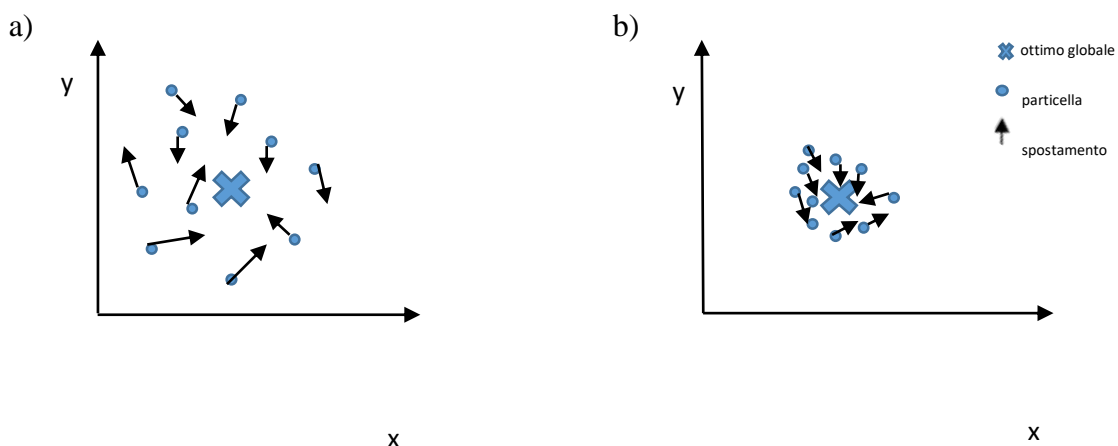


Figura 2.8: esempio del movimento delle particelle in uno spazio bidimensionale: nella a) si può vedere il movimento iniziale delle particelle che iniziano a spostarsi fino alla posizione ottima, come nella b).

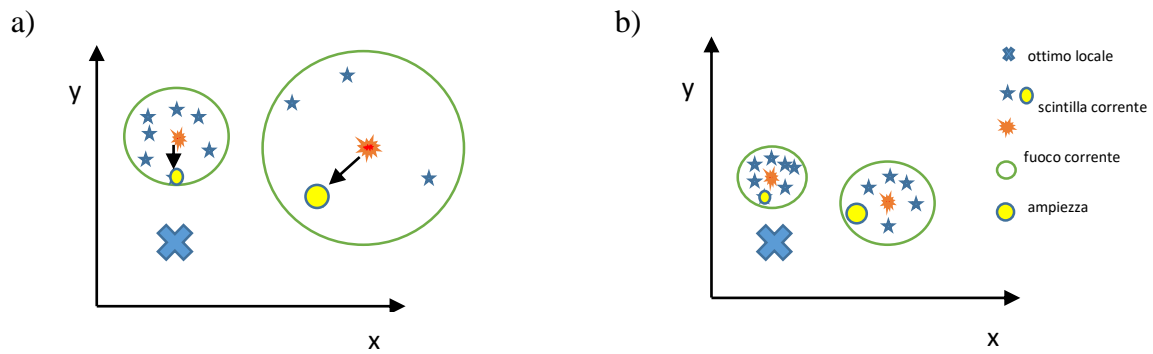


Figura 2.9: esempio bidimensionale del movimento delle scintille verso l'ottimo. Come si può nella figura a) avvengono le prime esplosioni. Successivamente vengono selezionate le scintille che saranno le posizioni dei nuovi fuochi d'artificio nell'iterazione successiva, come si può vedere dalla figura b).

## Capitolo 3

# MISURA DI RISCHIO COERENTE TWO-SIDED

*In questo capitolo verranno illustrate le principali critiche al modello di Markowitz. Per ovviare ai limiti del modello media-varianza si proporrà una nuova misura di rischio per il problema di selezione del portafoglio, la two-sided coherent risk measure. Inoltre s'introdurranno dei vincoli al problema di minimizzazione per renderlo più vicino alla realtà dei mercati finanziari.*

*In seguito, sarà riformulato il problema di selezione del portafoglio includendo sia la nuova misura di rischio che i vincoli. Il problema che si otterrà sarà un problema di minimizzazione vincolato; poiché le metaeuristiche, presentate nel capitolo 2, operano con problemi non vincolati, s'introdurrà il metodo della penalità esatta che permette di riformulare il problema vincolato in un problema non vincolato, attraverso la funzione di penalità.*

### 3.1 Critiche al modello media-varianza

Sebbene vada riconosciuto il merito a H. Markowitz di aver proposto un modello innovativo, questo non è esente da critiche; infatti come tutti i modelli non recepisce perfettamente tutti gli aspetti della realtà dei mercati finanziari, ponendo restrizioni e assunzioni troppo forti.

Di seguito si riportano le principali problematiche del modello di Markowitz.

- Non normalità. L'ipotesi sottostante il modello di Markowitz è che i rendimenti siano distribuiti come una distribuzione normale. Nella realtà le distribuzioni empiriche dei rendimenti deviano da quella che è la distribuzione normale. Facendo dei test sui diversi momenti delle distribuzioni empiriche dei rendimenti, studi come quello di Peirò (1999) e come quello di Scott e Hovarth (1980), hanno evidenziato come il valore delle statistiche differisca dal valore del corrispondente momento nella distribuzione normale.

Di seguito verranno analizzati due momenti della distribuzione empirica dei rendimenti, asimmetria e curtosi, evidenziando come questi, quando misurati su rendimenti reali, differiscano dai valori della normale.

I rendimenti sono simmetrici intorno alla media  $\mu$  se:

$$f(\mu + k) = f(\mu - k) \quad (3.1)$$

dove

$f$  è la funzione di densità di probabilità che rappresenta i rendimenti futuri,

$k$  è un numero reale qualsiasi.

Per verificare la presenza di asimmetria nei rendimenti, molti studi utilizzano l'indice di simmetria campionario e la sua distribuzione asintotica sotto l'ipotesi di normalità.

L'indice di simmetria campionario è così definito:

$$\hat{\alpha} = \frac{\sum_{t=1}^T \frac{(R_t - \bar{R})^3}{T}}{\hat{\sigma}^3} \quad (3.2)$$

dove

$T$  è il numero di osservazioni,

$R_t$  è il rendimento alla data  $t$ ,

$\hat{\sigma}$  è la deviazione standard campionaria,

$\bar{R}$  è la media campionaria.

La distribuzione normale presenta un indice di asimmetria pari a 0; invece la maggior parte delle serie dei rendimenti finanziari presenta un indice di asimmetria negativo, come evidenziato da Cassese e Pelagatti. L'asimmetria negativa nelle distribuzioni empiriche dal punto di vista economico indica che le perdite sono più importanti dei profitti.

La curtosi è l'allontanamento dalla distribuzione normale, rispetto alla quale si verifica o una maggior appiattimento (distribuzione platicurtica) o un maggior allungamento (distribuzione leptocurtica).

L'indice di curtosi campionario è definito come:

$$\hat{k} = \frac{\sum_{t=1}^T \frac{(R_t - \bar{R})^4}{T}}{\hat{\sigma}^4} \quad (3.3)$$

dove

$T$  è il numero di osservazioni,

$R_t$  è il rendimento alla data  $t$ ,

$\hat{\sigma}$  è la deviazione standard campionaria,

$\bar{R}$  è la media campionaria.

L'indice di curtosi campionario è pari a 3 nel caso di distribuzione normale. Se l'indice è maggiore di 3 si avrà una distribuzione cosiddetta leptocurtica, ossia la massa di probabilità associata alle realizzazioni delle code è maggiore rispetto a quella associata a una distribuzione normale; viceversa se l'indice è minore di 3, si parlerà di distribuzione platicurtica, ossia la massa delle probabilità associata alle realizzazioni delle code è minore rispetto a quella associata a una distribuzione normale. Diversi studi, tra cui quello di Peirò (1999), evidenziano come le distribuzioni empiriche siano affette da code spesse, ossia la distribuzione è leptocurtica.

Nello studio di Peirò viene anche considerato il Jarque-Bera test che analizza l'asimmetria e la curtosi congiuntamente. Più alto è il valore di questo test e più l'indice di asimmetria e l'indice di curtosi si discostano dai valori della distribuzione normale.

Anche in questo caso il Jarque-Bera test, nello studio di Peirò, assume valori elevati, portando a rifiutare la normalità delle distribuzioni empiriche dei rendimenti.

- Bassa diversificazione dei portafogli. Teoricamente, la strategia d'investimento razionale per un individuo avverso al rischio, che consideri solo i primi due momenti della distribuzione dei rendimenti, è un'opportuna diversificazione, poiché aggiungendo asset in portafoglio il rischio diminuisce. Tuttavia se si osserva il reale comportamento dei partecipanti al mercato si nota che la maggior parte degli investitori detiene portafogli non ottimamente diversificati. Nel 1975, come sottolineato nello studio di Blume, Marshall e Friend, la media degli asset detenuti in portafoglio, negli Stati Uniti, era meno di 4, e il 34% dei portafogli era formato da non più di due azioni. Questo comportamento, come evidenziano Conine e Tamarkin (1981), è dovuto alla preferenza degli investitori per un'asimmetria positiva della distribuzione dei rendimenti; i portafogli maggiormente diversificati presentano infatti dei rendimenti quasi simmetrici, mentre i portafogli poco diversificati presentano asimmetrie positive. Uno dei motivi per la preferenza dell'asimmetria positiva è che l'investitore non può perdere più del 100% del capitale investito, mentre i guadagni sono potenzialmente illimitati. Quindi la bassa diversificazione dei portafogli è giustificata dal fatto che aumentando il numero di asset nel portfolio, sebbene il rischio si riduca, aspetto positivo per l'investitore, si riduce anche l'asimmetria del portafoglio, aspetto invece considerato negativo per l'investitore.
- Funzione di utilità quadratica. Lo stesso Markowitz ha riconosciuto i limiti di questa funzione. La funzione di utilità è una parabola concava verso il basso ma con dominio limitato da 0 fino al punto di massimo; oltre il punto di massimo la parabola inizia a decrescere e l'assioma di non sazietà non è più rispettato. Se infatti il punto di massimo della funzione si trova all'interno del range dei rendimenti plausibili allora l'utilità marginale comincia a decrescere dopo il punto di massimo e i portafogli associati a quei rendimenti vengono considerati portafogli inferiori;
- Il modello di Markowitz presenta degli assunti alquanto irrealistici per un mercato reale, come l'assenza di costi di transizione e tasse, l'assenza di vincoli riguardanti le vendite allo scoperto e il numero di asset che può essere comprato o venduto.
- Varianza come misura di rischio. Uno dei principali limiti di questa misura è quello di trattare in egual modo sia i rendimenti positivi, il cosiddetto upside risk, che i rendimenti negativi, ossia il cosiddetto downside risk. Nella realtà, gli investitori trattano i guadagni

e le perdite in maniera diversa. Come affermano Scott e Hovarth (1980) il modello media-varianza può descrivere completamente il valore atteso della funzione di utilità quadratica solo se la distribuzione dei rendimenti è normale, uniforme o binomiale.

In merito a questi due ultimi limiti del modello di Markowitz, nel paragrafo successivo si propone una misura di rischio alternativa alla varianza, appartenente alla classe delle cosiddette misure di rischio coerenti, sia dei vincoli al problema di minimizzazione.

### 3.2 Misure di rischio coerenti

Le misure di rischio coerenti sono una classe di misure introdotte da Artzner et Al. (1998). Il lavoro di Artzner propone delle nuove misure di rischio da poter utilizzare per determinare il requisito di capitale extra per regolare il rischio assunto dagli investitori nel mercato. Se si considera  $X$  come la variabile casuale che indica il valore netto futuro dell'investimento, allora si definisce misura di rischio una funzione  $\rho$  che assegna un valore numerico al rischio di  $X$ . Quando  $\rho(X)$  è un numero positivo, questo sarà interpretato come la minima quantità di liquidità da aggiungere alla posizione rischiosa  $X$ , viceversa se è negativo, allora la quantità  $-\rho(X)$  sarà tolta dalla posizione. Uno dei vantaggi delle misure di rischio coerenti è quello di esprimere il rischio nella stessa misura del capitale investito, quindi in termini monetari, e non in termini percentuali.

Le misure di rischio coerenti per essere definite tali devono rispettare quattro assiomi, detti assiomi di coerenza:

- invarianza transizionale; se si investe parte del proprio capitale in un titolo privo di rischio, il rischio del portafoglio diminuisce tanto quanto è la quota investita nel titolo privo di rischio.

$$\rho(X + a \cdot r) = \rho(X) - \alpha \quad (3.4)$$

dove

$a$  è l'ammontare investito nella posizione priva di rischio,

$r$  è il rendimento dell'attività priva di rischio;

Questo assioma implica che per ogni  $X$ ,  $\rho(X + \rho(X)) = 0$ . In altri termini, aggiungendo una quantità priva di rischio pari a  $\rho(X)$  alla posizione  $X$  si ottiene una posizione priva di rischio.

- subaddittività. Date due posizioni  $X_1$  e  $X_2$ , allora:

$$\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2) \quad (3.5)$$

La subaddittività è alla base della diversificazione di portafoglio e implica che la combinazione di due posizioni rischiose dà un rischio minore rispetto alle due posizioni prese singolarmente.

- positività omogenea. Se l'investimento in una data posizione aumenta di una certa quantità,  $\lambda \geq 0$ , allora il rischio aumenta proporzionalmente all'investimento, ossia:

$$\rho(\lambda X) = \lambda \rho(X) \quad (3.6)$$

- monotonicità. Date due posizioni rischiose,  $X$  e  $Y$ , con  $X \preceq Y$ , allora il rischio associato alla posizione  $X$  sarà minore del rischio associato alla posizione  $Y$ ,

$$\rho(X) \leq \rho(Y) \quad (3.7)$$

In altre parole la monotonicità evidenzia la preferenza per un asset che fornisce sistematicamente rendimenti maggiori rispetto ad un altro asset.

L'assioma di monotonicità sembrerebbe contrastare con il principio della finanza secondo cui un maggior rendimento è accompagnato da un maggior rischio. Tuttavia questo postulato è vero solo quando si assume che nel mercato non vi siano possibilità di arbitraggio<sup>12</sup>. Se invece sono presenti, come spesso accade nella realtà dei mercati, la

---

<sup>12</sup> Un arbitraggio è un'operazione di acquisto e vendita della stessa attività su due mercati differenti, che permette di conseguire un guadagno certo senza correre alcun rischio. L'arbitraggio sfrutta le differenze di prezzo dello stesso bene quotato in due mercati differenti. Il differenziale di prezzo è dovuto all'inefficienza informativa o normativa. Ad esempio se nel mercato statunitense è negoziato un bene ad un minor prezzo rispetto al mercato italiano, l'arbitraggista può comprare nel mercato statunitense e rivendere nel mercato italiano, contribuendo inoltre a riallineare i due prezzi. Infatti l'aumento della domanda nel mercato statunitense fa aumentare il prezzo del bene, mentre l'aumento dell'offerta nel mercato italiano ne diminuisce il prezzo.

---

misura di rischio coerente che soddisfa l'assioma di monotonicità, aiuta a conseguire un guadagno risk-free. Infatti la misura di rischio coerente assegna al titolo meno remunerativo un rischio maggiore.

### 3.2.1 Two-sided coherent risk measure

Le misure di rischio coerenti two-sided, introdotte da introdotta da Chen e Wang (2008), sono diverse dalle maggioranza delle altra misure di rischio coerente poiché prendono in considerazione sia le deviazioni negative che positive dalla media.

Al fine di ottenere una misura di rischio che ben descriva la distribuzione dei rendimenti, la classe delle misure di rischio two-sided prende in considerazione i momenti di ordine superiore ottenuti attraverso la combinazione delle deviazioni positive e negative dalla media.

Nella classe di misure di rischio coerente, descritte nel paragrafo precedente, l'attenzione è posta o dal lato della domanda o dal lato dell'offerta. Questo implica che solo un lato della distribuzione è preso in considerazione nella misura di rischio associata. Tuttavia dal punto di vista della teoria dei giochi<sup>13</sup>, gli investitori dovrebbero prendere in considerazione l'azione del venditore di un asset quando vuole comprare quell'asset. In questo modo il comportamento rispetto a potenziali perdite è preso in considerazione sia dal lato della domanda che dal lato dell'offerta. In altre parole per ottenere una misura di rischio più idonea a prendere una solida decisione d'investimento, dovrebbe essere scelta una misura di rischio "two-sided".

Misure come quelle appartenenti alla classe delle misure di rischio nominate lower partial moment (LPM) o misure di rischio basate sul quantile, come il VaR (Value at Risk), operano una separazione tra downside risk e upside risk, non considerando parte dell'informazione. Tenendo conto di questo aspetto, ossia dell'uso completo di tutta l'informazione contenuta nella distribuzione di probabilità dei rendimenti, e preferendo che la misura di rischio rispetti gli assiomi di coerenza, la misura scelta per l'applicazione in questo elaborato come funzione obiettivo del problema di minimizzazione appartiene alla classe delle misure di rischio two-sided.

---

<sup>13</sup> La teoria dei giochi è una teoria matematica che analizza come un giocatore può massimizzare la propria ricchezza prendendo in considerazione le azioni e reazioni degli altri partecipanti al gioco.

### 3.2.2 Costruzione della nuova misura di rischio

La misura di rischio two-sided è costruita combinando sia il lato delle perdite che il lato dei profitti della variabile casuale rendimento.

Se si considera un solo periodo  $[0, D]$ , dove non è possibile alcuno scambio, allora il rischio può essere rappresentato dalla variabile casuale rendimento  $X$ , definita nello spazio di probabilità  $(\Omega, \mathcal{F}, \mathbb{P})$  di un asset o di un portafoglio. Se  $X$  è maggiore o uguale a zero si avrà una variabile casuale profitto, altrimenti se minore di zero si avrà una variabile casuale perdita. Misurare il rischio significa stabilire una corrispondenza  $\rho$  tra lo spazio di  $X$  e l'insieme dei numeri reali. Come nel caso delle misure di rischio coerenti, anche per le misure di rischio coerenti two-sided se  $\rho(X)$  è un numero positivo, questo sarà interpretato come la minima quantità di liquidità da aggiungere alla posizione rischiosa  $X$ , viceversa se è negativo, allora la quantità  $-\rho(X)$  sarà tolta dalla posizione.

Si definisce  $\|X\|_p = (E|X|^p)^{\frac{1}{p}}$ , dove  $E[\cdot]$  rappresenta il valore atteso della variabile casuale con  $p \in [1, +\infty]$ . Inoltre si definisce  $Y^- = \max(-Y, 0)$  e  $Y^+ = (-Y)^-$ . Infine si definisce  $\sigma_p^\pm(X) = \|(X - E[X])^\pm\|_p$ . Rispetto al valore atteso  $E[X]$ , la variabile casuale  $(X - E[X])^-$  rappresenta il lato delle perdite di  $X$ , mentre il la variabile casuale  $(X - E[X])^+$  rappresenta il lato dei profitti di  $X$ .  $\rho_{a,p}(X)$  rappresenta sia i momenti parziali superiori che inferiori di  $X$ , con  $\sigma_p^-(X)$  che rappresenta la “cattiva volatilità” e  $\sigma_p^+(X)$  che rappresenta la “buona volatilità”.

La misura di rischio coerente two-sided può quindi essere definita come:

$$\rho_{a,p} = \alpha \|(X - E[X])^+\|_1 + (1 - \alpha) \|(X - E[X])^-\|_p - E[X] \quad (3.8)$$

con  $p \in [1, \infty)$  e  $0 \leq \alpha \leq 1$

dove

$\alpha$  è una costante che bilancia tra “buona volatilità”,  $\sigma_1^+(X)$ , e “cattiva volatilità”, ovvero  $\sigma_p^-(X)$ ,

$p$  controlla il rischio dell'investimento. Più grande è  $p$  e più l'investitore è avverso al rischio.

La misura di rischio two-sided è quindi generata prendendo la combinazione tra la norma di indice 1 delle deviazioni positive e la norma di indice  $p$  delle deviazioni negative.

Dai valori assumibili dai parametri dai parametri  $a$  e  $p$  si può notare che  $\rho_{a,p}(X)$  è non decrescente rispetto a  $p$  e non-crescente rispetto ad  $a$ . Questo implica che più grande è il valore

assegnato a  $p$  più l'investitore è avverso al rischio, mentre più grande è il valore di  $a$ , meno l'investitore è avverso al rischio.

### 3.3 Applicazione della misura di rischio

L'applicazione della misura di rischio appena descritta sarà presentata in questo paragrafo. Prima però si specifica la seguente notazione che verrà utilizzata:

$x_i$  è il peso del titolo  $i$ -esimo contenuto nel portafoglio, con  $1 \leq i \leq n$ ;

$r_i$  è il tasso di rendimento del titolo  $i$ -esimo;

$\hat{r}_i = E[r_i]$  è il rendimento atteso dell' $i$ -esimo asset;

Il rendimento totale del portafoglio può quindi essere scritto come:

$$R = \sum_{i=1}^N x_i r_i. \quad (3.9)$$

Mentre il valore atteso del portafoglio sarà pari a:

$$E(R) = \hat{R} = \sum_{i=1}^N x_i \hat{r}_i. \quad (3.10)$$

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

$$\hat{r}_i = \left(\frac{1}{T}\right) \sum_{t=1}^T \hat{r}_{i,t}, \quad (3.11)$$

allora l'espressione campionaria dell'equazione (3.8) può essere riscritta come:

$$\rho_{a,p} = a \left\| \left( \sum_{i=1}^n x_i (r_i - \hat{r}_i) \right)^+ \right\|_1 + (1-a) \left\| \left( \sum_{i=1}^n x_i (r_i - \hat{r}_i) \right)^- \right\|_p - \hat{R} \quad (3.12)$$

Il primo termine dell'equazione (3.12) rappresenta le differenze positive tra la variabile  $R$  e il valore atteso della variabile, ossia  $E(R)$ , mentre il secondo termine rappresenta le differenze negative; la misura di rischio proposta è quindi una misura coerente 'two sided', poiché considera sia i momenti parziali superiori e sia quelli inferiori. La misura di rischio è quindi

generata prendendo la norma di indice 1 delle deviazioni positive, e la norma di indice  $p$  delle deviazioni negative. Applicando le due norme si ottiene:

$$\rho_{a,p}(R) = \frac{a}{T} \left[ \sum_{t=1}^T \left( \sum_{i=1}^N (r_{i,t} - \hat{r}_i) x_i \right)^+ \right] + (1-a) \left\{ \frac{1}{T} \sum_{t=1}^T \left[ \left( \sum_{i=1}^N (r_{i,t} - \hat{r}_i) x_i \right)^- \right]^p \right\}^{\frac{1}{p}} - \sum_{i=1}^n x_i \hat{r}_i. \quad (3.13)$$

### 3.4 Vincoli

In questo paragrafo vengono analizzati i vincoli che sono considerati all'interno del modello di selezione di portafoglio che si utilizzerà nell'applicazione. L'introduzione di questi vincoli permette di rendere il problema di selezione più vicino alla realtà dei mercati finanziari. Di seguito per completezza si riportano anche i due vincoli introdotti nel capitolo 1, ossia il vincolo di bilancio e il vincolo sul rendimento medio minimo.

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

$$\sum_{i=1}^N x_i = 1 \quad (3.14)$$

- vincolo sul rendimento medio minimo. Questo vincolo impone il livello di rendimento medio minimo che il portafoglio selezionato deve soddisfare, ossia:

$$\sum_{i=1}^N x_i \hat{r}_i \geq \pi \quad (3.15)$$

dove

$\hat{r}_i$  è il rendimento atteso dell' $i$ -esimo asset, come specificato nella (3.11).

Ponendo  $\hat{R} = \sum_{i=1}^N x_i \hat{r}_i$ , come indicato nella (3.10), il vincolo in (3.15) può essere riscritto come:

$$\hat{R} \geq \pi; \quad (3.16)$$

- vincolo di cardinalità. Questo vincolo è introdotto per contenere indirettamente i costi di transazione introducendo dei limiti sul numero di titoli da selezionare.

$$K_d \leq \sum_{i=1}^N z_i \leq K_u \quad \text{con } 1 \leq K_d \leq K_u \leq N \quad (3.17)$$

dove

$K_d$  è il numero minimo di titoli che il portafoglio potrà contenere,

$K_u$  è il numero massimo di titoli che il portafoglio potrà contenere,

$N$  è il numero di titoli,

$z_i$  è un vettore che assume valore 1 se il titolo  $i$ -esimo è stato selezionato, altrimenti assume valore 0.

Inoltre sarà necessario contenere le percentuali d'investimenti per ogni titolo, che saranno correlate con il minimo numero e con il massimo numero di titoli possibili nel portafoglio. S'introduce quindi un ulteriore vincolo:

$$z_i l_i \leq x_i \leq z_i u_i \quad \text{con } 0 \leq l_i \leq u_i \leq 1 \quad (3.18)$$

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.

Affinché i due vincoli appena illustrati siano compatibili bisogna imporre due disuguaglianze:

$$l_i \leq \frac{1}{K_l} \quad (3.19)$$

$$u_i \leq \frac{1}{K_u} \quad (3.20)$$

Il vincolo in (3.18) può essere visto come il vincolo di divieto di vendite allo scoperto; nei mercati, soprattutto nei momenti di crisi, spesso le vendite allo scoperto sono vietate per evitare che operare al ribasso degli investitori faccia crollare ancora di più i prezzi. Il divieto di vendite allo scoperto implica che le quote dei titoli non possano diventare negative; infatti nella (3.18) è espresso un vincolo di non negatività.

- il vincolo d'integralità, così formulato:

$$z_i(z_i - 1) = 0 \quad \text{con } i = 1, \dots, N \quad (3.21)$$

Il vincolo così formulato sostituisce la formulazione  $z_i \in \{0,1\}$ .

Il problema di selezione del portafoglio può così essere sintetizzato:

$$\begin{aligned} & \min_{x,z} \rho_{a,p}(R) & (3.22) \\ \text{s. t. } & \left\{ \begin{array}{l} \sum_{i=1}^N x_i = 1 \\ \hat{R} \geq \pi \\ K_d \leq \sum_{i=1}^N z_i \leq K_u \\ z_i d_i \leq x_i \leq z_i u_i \quad \forall i \\ z_i(z_i - 1) = 0 \quad \forall i \end{array} \right. \end{aligned}$$

### 3.5 Riformulazione del problema di selezione

Poichè la PSO e il FWA introdotti nel capitolo 2, sono metaeuristiche nate per problemi non vincolati, si riformula il problema vincolato in (3.22) in un problema non vincolato, utilizzando il metodo della penalità esatta (exact penalty method). Il metodo della penalità esatta prevede l'introduzione di una funzione di penalità, che penalizza la violazione dei vincoli, sommando quest'ultime in un'unica funzione di fitness che prende in considerazione sia la funzione obiettivo che i vincoli.

#### 3.5.1 Metodo delle penalità esatte

Sebbene esistano molti metodi per riformulare un problema vincolato in un problema non vincolato, si è scelto di utilizzare il metodo delle penalità esatte.

Il metodo consiste nel sostituire un problema generale di minimizzazione vincolato, formulato come:

$$\begin{aligned} & \min f(x) && (3.23) \\ \text{s. t. } & \begin{cases} h(x) = 0 \\ g(x) \leq 0 \end{cases} \end{aligned}$$

dove

$f(x)$  è una funzione continua in  $R^n$ ,

$h(x)$  è un vettore che ha come elementi  $m$  vincoli posti in forma d'uguaglianza,

$g(x)$  è un vettore che contiene  $p$  vincoli espressi in forma d'uguaglianza;

con un problema equivalente di minimizzazione non vincolato, formulato come:

$$\min f(x) + \varepsilon P(x) \quad (3.24)$$

dove

$\varepsilon$  è una costante positiva,

$P(x)$  è una funzione continua in  $R^n$  con le caratteristiche seguenti:

- $P(x) \geq 0 \quad \forall x \in R^n$ ,
- $P(x) = 0$  se e solo se  $x \in S$

Un aspetto fondamentale è capire quanto bene il problema non vincolato approssima il problema vincolato. Se si approssima il problema vincolato in un problema non vincolato attraverso la seguente funzione di penalità:

$$P(x) = \sum_{i=1}^m |h_i(x)| + \sum_{j=1}^p \max(0, g_j(x)) \quad (3.25)$$

allora il teorema delle penalità esatte garantisce la coincidenza tra il problema non vincolato con la funzione di penalità e il problema vincolato.

Il teorema delle penalità esatte ipotizza che se  $x^*$  è un punto che soddisfa le condizioni sufficienti del secondo ordine per essere un minimo locale del problema vincolato e se  $\omega$  e  $\mu$  sono due vettori aventi come elementi i moltiplicatori di Lagrange relativi rispettivamente agli  $m$  vincoli contenuti in  $h(x) = 0$  e ai  $p$  vincoli contenuti in  $g(x) \leq 0$ , allora per valori di  $\varepsilon > \max\{|\omega_i|, \mu_j; i = 1, \dots, m, j = 1, \dots, p\}$ ,  $x^*$  è un punto di minimo locale anche per il problema non vincolato con la funzione di penalità utilizzata in (3.25).

Per meglio comprendere il funzionamento del metodo della penalità esatta si riformulano i vincoli contenuti nel problema in (3.22) come segue:

$$\sum_{i=1}^N x_i = 1 \rightarrow \left| \sum_{i=1}^N x_i - 1 \right| = 0,$$

$$\hat{R} \geq \pi \rightarrow \max\{0, \pi - \hat{R}\} = 0,$$

$$K_d \leq \sum_{i=1}^N z_i \rightarrow \max\left\{0, K_d - \sum_{i=1}^N z_i\right\} = 0,$$

$$\sum_{i=1}^N z_i \leq K_u \rightarrow \max\left\{0, \sum_{i=1}^N z_i - K_u\right\} = 0,$$

$$z_i d_i \leq x_i \rightarrow \max\left\{0, \sum_{i=1}^N z_i - K_u\right\} = 0,$$

$$x_i \leq z_i u_i \rightarrow \sum_{i=1}^N \max\{0, x_i - z_i u_i\} = 0,$$

$$z_i \in \{0,1\} \rightarrow \sum_{i=1}^N |z_i(1 - z_i)|$$

Il problema riformulato in forma non vincolata diventa quindi:

$$\min_{X,Z,\varepsilon} P(X, Z, \varepsilon) \tag{3.26}$$

dove

$$P(X, Z, \varepsilon) = p_{a,p}(R) + \frac{1}{\varepsilon} [|\sum_{i=1}^N x_i - 1| + \max\{0, \pi - \hat{R}\} + \max\{0, K_d - \sum_{i=1}^N z_i\} + \max\{0, \sum_{i=1}^N z_i - K_u\} + \sum_{i=1}^N \max\{0, z_i d - x_i\} + \sum_{i=1}^N \max\{0, x_i - z_i u\} + \sum_{i=1}^N |z_i(1 - z_i)|] \tag{3.27}$$

Analizzando la (3.27), si può notare come alla misura di rischio vengono sommati tutti i vincoli, corretti per  $\frac{1}{\varepsilon}$ .

Se ad esempio si prende il vincolo di bilancio riformulato, si può affermare che se la somma delle quote degli  $N$  titoli che compongono uno dei possibili portafogli è minore di 1, ossia se  $\sum_{i=1}^N x_i < 1$ , il vincolo di bilancio non è rispettato. Di conseguenza la quantità  $\sum_{i=1}^N x_i - 1$  sarà minore di zero. Per evitare che la misura di rischio diminuisca, viene applicato il valore assoluto, ossia  $|\sum_{i=1}^N x_i - 1|$ . Questa violazione del vincolo viene moltiplicata per  $\frac{1}{\varepsilon}$  e sommata alla misura di rischio. Se invece il vincolo è rispettato allora  $\sum_{i=1}^N x_i = 1$ ; sostituendo si avrà  $|1 - 1| = 0$ . Questo implica che alla misura di rischio non verrà sommata alcuna quantità. In un problema di selezione di portafoglio dove l'obiettivo è quello di minimizzare la funzione di fitness, è chiaro che più uno dei possibili portafogli non rispetta i vincoli, più la funzione di fitness assume un valore elevato, portando gli algoritmi a scartare tale portafoglio.

## Capitolo 4

# APPLICAZIONE DELLE METAEURISTICHE

*In questo capitolo conclusivo si propone l'applicazione delle due metaeuristiche alle serie storiche dei prezzi di 30 titoli del mercato statunitense. In particolare vengono analizzati tre scenari, corrispondenti ad altrettanti periodi economici, per verificare come il modello si comporti in situazioni di mercato differenti.*

### 4.1 Informazioni preliminari

In questo capitolo si applica il modello di selezione del portafoglio a dati reali con lo scopo di testarne la capacità. Entrambi gli algoritmi della PSO e del FWA sono stati processati in MATLAB 2016.

L'analisi è svolta utilizzando le serie storiche di 30 titoli del mercato statunitensi, scelti tra i 100 titoli più attivi al 30/09/2017. Di seguito sono riportate le emittenti dei titoli selezionati, con i relativi codici della borsa: Rite Aid Corporation (RAD), Bank of America (BAC), Blackberry (BBRY), Oracle Corporation (ORLC), E.I. Du Pont de Nemours and Company (DD), Ambev (ABEV), Advanced Micro Devices (AMD), Micron Technology (MU), General Electric Company (GE), Ford Motor Company (F), Vale (VALE), Apple (AAPL), Chesapeake Energy Corporation (CHK), Enterprise Products Partners L.P. (EPD), Applied Materials (AMAT), Intel Corporation (INTC), Wells Fargo & Company (WFC), Transocean LTD. (RIG), AT&T (T), Astoria Financial Corporation (AF), Comcast Corporation (CMCSA), Brocade Communications Systems (BRCD), KB-Home (KBH), Sirius XM Holdings (SIRI),

Verizon Communications (VZ), Marathon Oil Corporation (MRO), Citigroup (C), Encana Corporation (ECA), The Kroger Co. (KR), Valbant Pharmaceuticals International (VRX).

Dai prezzi delle serie storiche dei titoli sopra elencati, sono stati ricavati i rendimenti logaritmici come segue:

$$R_{i,t} = \ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right) \quad (4.1)$$

dove,

$R_{i,t}$  è il rendimento al tempo  $t$  dell' $i$ -esimo titolo,

$P_{i,t}$  è il prezzo al tempo  $t$  dell' $i$ -esimo titolo,

$P_{i,t-1}$  è il prezzo al tempo  $t-1$  dell' $i$ -esimo titolo.

La scelta di utilizzare il rendimento logaritmico, al posto di quello percentuale, è data dalla supposizione che i prezzi siano generati dalla legge finanziaria dell'interesse continuo<sup>14</sup>. Tuttavia poiché le osservazioni dei prezzi prese sono giornaliere, l'effettiva differenza tra il rendimento percentuale e il rendimento logaritmo è trascurabile.

Il procedimento si articola in due passi:

- individuazione di alcuni scenari economici per l'analisi del modello; questa suddivisione è giustificata dal fatto che non avrebbe senso fare un'analisi sugli ultimi 10 anni, visto gli importanti eventi economici;
- la divisione dei rendimenti per ogni scenario in due parti: in-sample e out-of-sample.

La parte in-sample comprende i primi 12 mesi, mentre la parte out-of-sample gli ultimi 6 mesi. L'analisi viene svolta ipotizzando di essere alla fine del dodicesimo mese e di voler investire il proprio capitale per i successivi 6 mesi. Se i rendimenti e le misure di rischio del periodo in-sample sono simili al periodo out-of-sample allora il modello dovrebbe suggerire le migliori percentuali da investire nei titoli selezionati dal modello stesso.

Per quanto riguarda invece gli scenari si è deciso di scegliere tre periodi differenti dell'economia statunitense:

---

<sup>14</sup> Poiché la quotazione avviene nel continuo, la legge finanziaria da utilizzare dovrebbe essere la seguente:  $P_t = P_{t-1}e^{r_{i,t}}$ , da cui si ricava il rendimento logaritmico calcolato nella 4.1.

- lo scenario n. 1 è compreso tra il gennaio 2008 e il giugno 2009. Questo periodo è particolarmente significativo per l'economia statunitense in quanto comprende la crisi del mercato immobiliare e il fallimento di Lehman Brothers.
- lo scenario n. 2 è compreso tra il gennaio 2012 e il giugno 2013. Nel 2012 l'economia statunitense è stata deludente, e sotto le stime d'inizio anno; il PIL si è attestato all'1,7% e la disoccupazione è rimasta su livelli elevati; si è registrato un aumento del debito pubblico in percentuale del PIL. Invece per quanto riguarda il periodo out-of-sample, nel 2013 gli Stati Uniti hanno registrato una ripresa economica.
- lo scenario n. 3 è compreso tra il gennaio 2016 e giugno 2017. Nel 2016 l'economia statunitense è cresciuta meno del previsto, non accelerando come molti investitori si attendevano, trend che si sta confermando anche per il 2017.

Prima di procedere con l'analisi appena descritta si conducono alcuni test preliminari per la PSO e il FWA.

## 4.2 Parametri

Prima di applicare il modello ai diversi scenari è necessario settare i parametri riguardanti sia i vincoli e la misura di rischio e sia i parametri della PSO e del FWA.

Per quanto riguarda i parametri e i vincoli del modello in ogni prova effettuata sono stati mantenuti alcuni valori stabili:

- numero minimo di titoli detenibili in portafoglio:  $K_d=7$ ;
- numero massimo di titoli detenibili in portafoglio:  $K_u=28$ ;
- percentuale d'investimento minima in ogni titolo:  $d_i=3\% \forall i$ ;
- percentuale d'investimento massima in ogni titolo:  $u_i=20\% \forall i$ ;
- parametri della misura di rischio:  $a=0.5$  e  $p=2$ .
- il rendimento obiettivo fissato pari alla media dei rendimenti attesi in-sample di ogni titolo:

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

Il rendimento medio giornaliero desiderato dipende dalle serie storiche dei prezzi di ogni scenario, quindi i valori saranno diversi per ciascuno dei tre scenari.

#### 4.2.1 PSO-impostazione dei parametri

Per quanto riguarda la PSO si è deciso di scegliere la versione con il peso d'inerzia. I parametri che calcolano la velocità di aggiornamento di ogni particella sono stati settati come indicato in letteratura, ossia:

- peso d'inerzia:  $w=0.7298$ ;
- coefficienti di accelerazione cognitivo:  $c1=1.49618$ ;
- coefficiente di accelerazione sociale:  $c2=1.49618$ ;
- numero di particelle: 50;
- numero d'iterazioni: 1000.

L'ultimo parametro da settare è il parametro della funzione di penalità  $\epsilon$ . È un parametro di notevole rilievo in quanto garantisce la corrispondenza tra il problema originario vincolato, e il problema non vincolato. Il corretto settaggio è dunque di fondamentale importanza. Si è scelto di effettuare 5 prove per sette diversi valori di  $\epsilon$  con una popolazione pari a 50 e un numero d'iterazioni pari a 1000. Nella Tabella 4.1 sono riportati i valori medi standardizzati e la relativa deviazione standard per ogni valore di  $\epsilon$ .

$\epsilon$	Fitness Normalizzata	Deviazione Standard
1	0.555864	0.336653
0,1	0.600245	0.305207
0,01	0.355839	0.237961
0,001	0.301886	0.299777
0,0001	0.253714	0.196157
0,00001	0,226385	0,071127
0,000001	0,443365	0.278555
0,0000001	0,183001	0,084725

Tabella 4.1: settaggio del parametro  $\epsilon$

Dalla Tabella 4.1 si può vedere che la fitness normalizzata con il valore più basso è quella corrispondente a  $\epsilon = 0,0000001$

### 4.2.2 FWA-impostazione dei parametri

Per quanto riguarda invece il FWA i parametri sono stati così settati, come indicato in letteratura:

- numero fuochi d'artificio:  $N=5$ ;
- l'ampiezza massima di ogni fuoco d'artificio:  $A_{\max}=40$ ;
- il parametro per il numero massimo di scintille è pari a 0.8;
- il parametro per il numero minimo di scintille è pari a 0.04;
- il limite massimo dello spazio:  $x_{\max}=100$ ;
- il limite minimo dello spazio:  $x_{\min}=-100$ ;
- parametro che controlla il numero totale di scintille:  $m=40$ ;
- numero di scintille prodotto dall'esplosione Gaussiana per ogni fuoco d'artificio:  $m_{\text{gauss}}=1$ ;
- numero di iterazione:  $k=1000$ ;

Oltre a questa configurazione si propone anche una seconda configurazione che prevede un numero un numero d'iterazioni pari a 10000 e un numero di scintille pari a 10.

### 4.3 Applicazione

Dopo aver settato i parametri, si testa il modello di selezione del portafoglio per diversi valori di  $a$  e  $p$  e per diversi valori di  $K_u$ , considerando un anno dei rendimenti giornalieri per i tre diversi scenari. Questo test è svolto per verificare la proprietà di non-crescenza di  $a$  rispetto a  $\rho_{a,p}$  e la proprietà di non-decrescenza di  $p$  rispetto a  $\rho_{a,p}$ .

Il numero massimo d'iterazione è stato settato a 1000 per la PSO e il FWA mentre il numero massimo di titoli da detenere in portafoglio è stato settato pari a 10 e a 30.

Nella tabella 4.2 sono stati calcolati, per la PSO, i valori di  $\rho_{a,p}$  fissando  $a$  pari a 0,5 e facendo variare  $p$ . Come si può vedere all'aumentare di  $p$  aumenta anche la misura di rischio, soddisfacendo così la proprietà di monotonia di  $\rho_{a,p}$ .

<i>Scenario 1</i>	$p=1$	$p=2$	$p=3$
$\rho_{0,5,p}: K_u = 10$	0.003536	0.004781	0.014793
Titoli selezionati	9	10	9

$\rho_{0,5,p}: K_u = 30$	0.005127	0.005426	0.005484
Titoli selezionati	22	17	13
<hr/>			
<i>Scenario 2</i>	<i>p=1</i>	<i>p=2</i>	<i>p=3</i>
$\rho_{0,5,p}: K_u = 10$	0.002927	0.003006	0.0033315
Titoli selezionati	8	8	6
<hr/>			
$\rho_{0,5,p}: K_u = 30$	0.000183	0.007396	0.008421
Titoli selezionati	13	15	17
<hr/>			
<i>Scenario 3</i>	<i>p=1</i>	<i>p=2</i>	<i>p=3</i>
$\rho_{0,5,p}: K_u = 10$	0.000661	0.006001	0.009191
Titoli selezionati	8	8	7
<hr/>			
$\rho_{0,5,p}: K_u = 30$	0.002696	0.0070684	0.007836
Titoli selezionati	11	17	18

Tabella 4.2: PSO-monotonia di  $\rho_{a,p}$  per  $a=0.5$  con diversi valori di  $p$

Nella tabella 4.3 sono invece riportati i valori di  $\rho_{a,p}$  per  $p=2$ . Come si può notare, anche in questo caso la proprietà di monotonia di  $\rho_{a,p}$  è soddisfatta. La proprietà di monotonia può essere utilizzata per riflettere il comportamento di un individuo rispetto al rischio. La proprietà di non decrescenza di  $a$  rispetto a  $\rho_{a,p}$  significa che  $\rho_{a,p}$  con un valore elevato di  $a$  è collegato a un investitore meno avverso al rischio. Mentre più alto è il valore di  $p$  e più l'investitore associa a quell'investimento un valore della misura di rischio più elevato.

<i>Scenario 1</i>	<i>a=0</i>	<i>a=0.5</i>	<i>a=1</i>
$\rho_{a,2}: K_u = 10$	0.006269	0.006261	0.004061
Titoli selezionati	8	9	7
<hr/>			
$\rho_{a,2}: K_u = 28$	0.006856	0.006762	0.003495
Titoli selezionati	17	16	21
<hr/>			
<i>Scenario 2</i>	<i>a=0</i>	<i>a=0.5</i>	<i>a=1</i>
$\rho_{a,2}: K_u = 10$	0.002681	0.001040	0.001040
Titoli selezionati	9	9	8
<hr/>			
$\rho_{a,2}: K_u = 28$	0.004784	0.001071	0.000008

Titoli selezionati	17	12	14
<i>Scenario 3</i>	$a=0$	$a=0.5$	$a=1$
$\rho_{0,5,p}: K_u = 10$	0.015527	0.003870	0.003137
Titoli selezionati	10	9	7
$\rho_{0,5,p}: K_u = 28$	0.000884	0.000827	0.000394
Titoli selezionati	12	17	13

Tabella 4.3. PSO-monotonia di  $\rho_{a,p}$  per  $p=2$  con diversi valori di  $a$

Di seguito si riportano anche le tabelle riguardanti il FWA. Come si può vedere anche in questo caso la proprietà di monotonia è soddisfatta. Inoltre si può notare come nel primo periodo i valori della misura di rischio aumentino in modo consistente con l'aumentare di  $K_u$ .

Tabella 4.4: PSO-monotonia di  $\rho_{a,p}$  per  $a=0.5$  con diversi valori di  $p$

<i>Scenario 1</i>	$p=1$	$p=2$	$p=3$
$\rho_{0,5,p}: K_u = 10$	0.005834	0.009254	0.016668
Titoli selezionati	4	4	5
$\rho_{0,5,p}: K_u = 28$	0.104322	0.150761	0.174606
Titoli selezionati	5	5	5
<i>Scenario 2</i>	$p=1$	$p=2$	$p=3$
$\rho_{0,5,p}: K_u = 10$	0.029751	0.047432	0.0556063
Titoli selezionati	5	5	5
$\rho_{0,5,p}: K_u = 28$	0.036847	0.050403	0.112967
Titoli selezionati	5	5	6
<i>Scenario 3</i>	$p=1$	$p=2$	$p=3$
$\rho_{0,5,p}: K_u = 10$	0.050643	0.059236	0.067311
Titoli selezionati	4	5	5
$\rho_{0,5,p}: K_u = 28$	0.050327	0.056083	0.056351
Titoli selezionati	4	4	6

Tabella 4.5: FWA-monotonia di  $\rho_{a,p}$  per  $p=2$  con diversi valori di  $a$ 

<i>Scenario 1</i>	$a=0$	$a=0.5$	$a=1$
$\rho_{a,2}: K_u = 10$	0.210326	0.173355	0.103642
Titoli selezionati	5	5	4
$\rho_{a,2}: K_u = 28$	0.189238	0.174492	0.134937
Titoli selezionati	6	5	5
<i>Scenario 2</i>	$a=0$	$a=0.5$	$a=1$
$\rho_{a,2}: K_u = 10$	0.050995	0.045523	0.040703
Titoli selezionati	5	5	5
$\rho_{a,2}: K_u = 28$	0.056092	0.049602	0.027740
Titoli selezionati	5	5	4
<i>Scenario 3</i>	$a=0$	$a=0.5$	$a=1$
$\rho_{0,5,p}: K_u = 10$	0.068265	0.049972	0.026856
Titoli selezionati	5	5	5
$\rho_{0,5,p}: K_u = 28$	0.075215	0.042206	0.039132
Titoli selezionati	5	5	5

Dal test appena effettuato si può notare come nel caso della PSO i portafogli sono più diversificati rispetto a quelli del FWA. Infatti i portafogli del FWA hanno in media un numero di titoli pari a 5, che è il valore del parametro  $K_d$  settato per il test sulla verifica della proprietà di monotonia. Il FWA tende quindi a selezionare un numero di titoli pari al numero minimo di titoli detenibile in portafoglio.

### 4.3.1 Scenari

In questo paragrafo si confrontano i due algoritmi per i tre diversi scenari, soffermandosi in particolare sulle violazioni dei vincoli.

#### 4.3.1.1 Scenario 1

Il primo scenario va dal 01/01/2008 e al 30/06/2009 e comprende 230 osservazioni di rendimenti giornalieri *in-sample* e 125 rendimenti giornalieri *out-of-sample*. I parametri sono

stimati in un periodo che comprende sia la crescita economica, ma anche l'inizio della crisi del mercato immobiliare statunitense e dei mutui subprime con il conseguente fallimento di Lehman Brothers. La valutazione del modello viene fatta nei primi sei mesi del 2009, quindi in un contesto di piena crisi. Il rendimento giornaliero desiderato per questo periodo è pari alla media dei rendimenti in-sample, ossia  $\pi = -0.00319384085399698$ . Si ricorda che per il FWA sono state utilizzate due configurazioni differenti per analizzare il comportamento dell' algoritmo al variare del numero delle scintille e al variare del numero d' iterazioni.

La prima configurazione prevede un numero di scintille pari a 10 con 10000 iterazioni, mentre la seconda configurazione prevede un numero di scintille pari a 40 con 1000 iterazioni.

	Configurazione 1	Configurazione 2
<i>fitness</i>	27802049	41148950
<i>Misura di rischio out of sample</i>	0.0302327128759548	0.0257338869543547
<i>Misura di rischio in sample</i>	0.167891188965218	0.136906085783197
<i>Rendimento out of sample</i>	0.000222017485116557	0.00350283161150273
<i>Rendimento in sample</i>	-0.00288643224030923	-0.00488264483123915
<i>Vincolo di bilancio</i>	0.0277036711459553	0.0300235543537921
<i>Vincolo rendimento atteso</i>	0	0.00168880397724217
<i>Vincolo numero minimo di titoli</i>	0.0123359545013582	2.14671819331047
<i>Vincolo numero massimo di titoli</i>	0	0
<i>Vincolo quota minima di titoli</i>	0.540258726735364	0.291816793547645
<i>Vincolo quota massima di titoli</i>	0.0166771328916919	0.0227438776149755
<i>Vincolo <math>z \in \{0;1\}</math></i>	2.18322948530627	1.62190383377072
<i>Vincolo di non negatività</i>	0	0
<i>numero titoli scelti</i>	6.98766404549864	4.85328180668953
<i>Somma quote ottime</i>	1.02770367114596	1.030023554353793

Tabella 4.6. Output scenario 1 FWA

Da una prima analisi dei dati si può notare come la misura di rischio out-of-sample sia notevolmente inferiore rispetto al periodo in-sample per entrambe le configurazioni, mentre il rendimento risulta leggermente superiore nel periodo in-sample. Per quanto riguarda invece i vincoli si può notare che entrambe le configurazioni non riescono a rispettare il vincolo d'integralità e il vincolo della quota minima di titoli presenti in portafoglio. Questa seconda violazione può essere giustificata dal fatto che nel 2008-2009 i prezzi hanno subito un forte calo: una soluzione potrebbe essere quella delle vendite allo scoperto, ma il modello non contempla tale ipotesi in quanto spesso nei mercati finanziarie le vendite allo scoperto sono vietate, soprattutto nei periodi di crisi. La somma delle quote ottime è circa 1; tuttavia bisogna tenere conto del fatto che il vincolo d'integralità è stato violato; per ottenere il rendimento riportato in tabella bisognerebbe quindi poter comprare una percentuale del titolo, cosa nella realtà non possibile. Si nota inoltre che il vincolo del numero minimo di titoli detenibile nel portafoglio non è stato rispettato nella seconda configurazione con un numero maggiore di scintille. Nei vari test si è infatti notato che un aumento del numero di scintille porta il vettore  $z \in \{0;1\}$  a propendere per dei valori pari a 0 più che a 1, comportando l'esclusione di titoli oltre il limite consentito dal valore di  $K_d$ . Sembrerebbe quindi che un aumento della numerosità delle scintille porti a un portafoglio meno diversificato rispetto a un portafoglio dove il numero di scintille è limitato a un valore più basso. Questa ipotesi è confermata anche dal valore della violazione del vincolo d'integralità, leggermente più bassa nella seconda configurazione, visto la presenza di molti valori del vettore  $z$  prossimi a zero.

Per quanto riguarda invece la PSO si può notare dalla Tabella 4.7 che i valori del rendimento ma anche la misura di rischio sono più elevati nel periodo out-of-sample. A differenza del FWA i vincoli presentano delle violazioni trascurabili. Inoltre come già accennato il portafoglio della PSO presenta un maggior grado di diversificazione rispetto al FWA.

<i>fitness</i>	38662
<i>Misura di rischio out of sample</i>	0.0118748311433567
<i>Misura di rischio in sample</i>	0.00484804300702833
<i>Rendimento out of sample</i>	0.00144993812947637
<i>Rendimento in sample</i>	-0.00261375920834987
<i>Vincolo di bilancio</i>	0.000209778947837913
<i>Vincolo rendimento atteso</i>	0

<i>Vincolo numero minimo di titoli</i>	0
<i>Vincolo numero massimo di titoli</i>	0
<i>Vincolo quota minima di titoli</i>	3.86081831487454e-06
<i>Vincolo quota massima di titoli</i>	5.74203187157801e-05
<i>Vincolo <math>z \in \{0;1\}</math></i>	0.00381229570789406
<i>Vincolo di non negatività</i>	0
<i>numero titoli scelti</i>	19
<i>Somma quote ottime</i>	1.00000940780985

Tabella 4.7: PSO scenario 2

#### 4.2.2.2 Scenario 2

Il secondo scenario è compreso tra il 01/01/2012 e il 30/06/2013. Il periodo in-sample è formato dai primi 12 mesi, nello specifico dalle prime 228 osservazioni, mentre il secondo dalle rimanenti 125 osservazioni relative al primo semestre del 2013. Nel 2012 l'economia statunitense è stata ben al di sotto delle aspettative stimate a inizio anno, mentre il 2013 ha visto una lieve ripresa economica.

Anche in questo secondo scenario il FWA riesce a ottenere dei risultati leggermente migliori rispetto alla PSO per quanto riguarda i rendimenti e la misura di rischio out-of-sample. Tuttavia bisogna prendere in considerazione anche le violazioni dei vincoli; infatti il valore della fitness, con lo stesso settaggio di  $\epsilon$ , risulta di gran lunga maggiore nel FWA.

	Configurazione 1	Configurazione 2
<i>fitness</i>	36720580	40510081
<i>Misura di rischio out of sample</i>	0.00520260111371553	0.00462305792874613
<i>Misura di rischio in sample</i>	0.0699738686993903	0.0831872856129512
<i>Rendimento out of sample</i>	0.00123193355626729	0.00105344558397417
<i>Rendimento in sample</i>	-7.73209815299725e-05	-0.00192408743854748

<i>Vincolo di bilancio</i>	0.0712786707283746	0.0718176987691975
<i>Vincolo rendimento atteso</i>	0.00032847933240188	0.00217524578941939
<i>Vincolo numero minimo di titoli</i>	0	0
<i>Vincolo numero massimo di titoli</i>	0	0
<i>Vincolo quota minima di titoli</i>	0.214313738750465	1.59939985652754
<i>Vincolo quota massima di titoli</i>	0.463281961575595	0.163414510764468
<i>Vincolo <math>z \in \{0;1\}</math></i>	2.92285521572965	2.2142008739828
<i>Vincolo di non negatività</i>	0	0
<i>numero titoli scelti</i>	7.97667881276372	8.0842775733648
<i>Somma quote ottime</i>	1.02770300867751	0.928182301230802

Tabella 4.8. Scenario 2 FWA

<i>fitness</i>	615152
<i>Misura di rischio out of sample</i>	0.00358444657249427
<i>Misura di rischio in sample</i>	0.000826293404620552
<i>Rendimento out of sample</i>	0.000611499191894129
<i>Rendimento in sample</i>	0.00123661711468242
<i>Vincolo di bilancio</i>	0.000753204858050527
<i>Vincolo rendimento atteso</i>	0
<i>Vincolo numero minimo di titoli</i>	0
<i>Vincolo numero massimo di titoli</i>	0
<i>Vincolo quota minima di titoli</i>	0.0599397407077914
<i>Vincolo quota massima di titoli</i>	5.61559330656747e-05
<i>Vincolo <math>z \in \{0;1\}</math></i>	0.00147417711310553

<i>Vincolo di non negatività</i>	2.89193690429596e-08
<i>numero titoli scelti</i>	17.999784012595
<i>Somma quote ottime</i>	0.999980750733844

Tabella 4.9. Scenario 2 -PSO

#### 4.2.2.3 Scenario 3

Il terzo scenario include 357 osservazioni dal 01/01/2016 al 30/06/2017, di cui 230 osservazioni *in-sample* e 125 *out-of-sample*. Nel 2016 l'economia statunitense non ha avuto l'accelerazione che molti investitori si aspettavano e anche il primo semestre del 2017 ha fatto registrare un trend simile. Dai dati della Tabella 4.3 si può notare come il rischio sia sceso nel periodo *out-of-sample*; tuttavia l'abbassamento del rischio è seguito anche da un notevole calo del rendimento nel periodo *out-of-sample*. Il rendimento desiderato è pari a 0.000586776195118273.

In questo scenario si può notare come la configurazione 1 del FWA abbia una violazione del vincolo della quota minima trascurabile, aspetto finora non osservato negli altri due scenari. Inoltre si verifica un aumento della violazione del vincolo della quota massima detenibile per ogni titolo rispetto ai due precedenti scenari. Anche la PSO registra un aumento, seppur molto contenuto, della violazione del vincolo della quota massima di ogni titolo, accompagnato da una diminuzione nel numero di titoli detenuti in portafoglio, forse indice di una migliore condizione economica rispetto a quella del primo scenario.

	Configurazione 1	Configurazione 2
<i>fitness</i>	31300567	36221431
<i>Misura di rischio out of sample</i>	0.0058666427662899	0.00489978293249838
<i>Misura di rischio in sample</i>	0.0710183830067211	0.048286238935659
<i>Rendimento out of sample</i>	0.000291970485402003	0.000461412391081723
<i>Rendimento in sample</i>	0.00103602663748725	0.000991200492515379
<i>Vincolo di bilancio</i>	0.0468825551051704	0.00956714799655112

<i>Vincolo rendimento atteso</i>	0	0
<i>Vincolo numero minimo di titoli</i>	0.529381760032956	0.698478135348183
<i>Vincolo numero massimo di titoli</i>	0	0
<i>Vincolo quota minima di titoli</i>	0.0573367781908117	0.440864848577011
<i>Vincolo quota massima di titoli</i>	0.143707186226273	0.141811954725604
<i>Vincolo <math>z \in \{0;1\}</math></i>	2.35274843890846	2.33142108268028
<i>Vincolo di non negatività</i>	0	0
<i>numero titoli scelti</i>	6.47061823996704	6.30152186465182
<i>Somma quote ottime</i>	0.95311744489483	1.009407861702490

Tabella 4.10. Scenario 3 FWA

<i>fitness</i>	253300
<i>Misura di rischio out of sample</i>	0.00337438364204777
<i>Misura di rischio in sample</i>	0.000209916554886023
<i>Rendimento out of sample</i>	0.000264929505379455
<i>Rendimento in sample</i>	0.000630974810225472
<i>Vincolo di bilancio</i>	0.00484210311024746
<i>Vincolo rendimento atteso</i>	0
<i>Vincolo numero minimo di titoli</i>	0
<i>Vincolo numero massimo di titoli</i>	0
<i>Vincolo quota minima di titoli</i>	0.00073567268821569
<i>Vincolo quota massima di titoli</i>	0.00122751184697997
<i>Vincolo <math>z \in \{0;1\}</math></i>	0.0235434562612161
<i>Vincolo di non negatività</i>	4.86115987814776e-41

---

<i>numero titoli scelti</i>	17.0163712264715
<i>Somma quote ottime</i>	0.999466440321013

*Tabella 4.11. Scenario 3 PSO*

## **Conclusione**

Nel seguente elaborato, sono state applicati due differenti algoritmi al problema di selezione del portafoglio con la misura di rischio coerente two-sided. Nell'ultimo capitolo si sono confrontati i risultati ottenuti dalla PSO e dal FWA sui rendimenti di 30 titoli del mercato statunitense. Sebbene il FWA fornisce soluzioni con portafogli aventi valori dei rendimenti maggiore rispetto ai valori della PSO, e valori della misura di rischio minori o uguali, si può concludere che il FWA non riesce a trovare soluzioni che rispettano i vincoli imposti: in particolare non si hanno violazioni per quanto riguarda il vincolo di non negatività, il vincolo di bilancio e il vincolo di redditività, ma si ha una forte violazione del vincolo di integralità. L'algoritmo inoltre predilige una bassa diversificazione del portafoglio rispetto alla PSO. Si può quindi affermare che l'operazione di spostamento, il numero delle scintille e l'ampiezza del fuoco d'artificio del FWA che dovrebbero condurre verso un "buon fuoco d'artificio", ossia verso una buona soluzione, non sono misure abbastanza efficienti come invece lo è la velocità calcolata nella PSO che determina lo spostamento di ogni particella. Inoltre l'utilizzo della distanza nel FWA per impedire che l'algoritmo degeneri a una soluzione troppo velocemente, forse non garantisce la selezione delle scintille con le migliori fitness. Una modifica apportabile al FWA potrebbe essere proprio il criterio di selezione delle scintille che costituiscono la nuova popolazione di fuochi d'artificio nell'iterazione successiva, magari selezionando più di una sola scintilla secondo il criterio della fitness migliore.

## Appendice A

Quote ottime d'investimento in ciascuno dei 30 titoli componenti il portafoglio per ciascuno dei 3 scenari.

FWA quote ottime: scenario 1		
	Configurazione 1	Configurazione 2
Rite Aid Corporation	0.00766519415494188	4.45410477153382e-06
Bank of America	1.06251032445096e-08	0.00221968053857154
Blackberry	6.5257365993003e-06	0.00325402033676735
Oracle Corporation	0.000650599316404956	1.95808600424764e-08
E.I. Du Pont de Nemours and Company	0.0475223546932314	1.103808582549e-08
Ambev	0.000367204510663091	0.0380184446156967
Advanced Micro Devices	9.50853441100836e-07	0.0857467700961463
Micron Technology	2.53062926858642e-05	0.254776807551212
General Electric Company	1.16048512959629e-09	2.54367726832195e-06
Ford Motor Company	0.00535355935906336	0.158739065280233
Vale	0.0575739122556817	0.0462547798671087
Apple	0.00055603077971897	0.0753849199424693
Chesapeake Energy Corporation	0.144653823138131	0.0321138383337359
Entreprise Products Partners L.P.	2.05272093504388e-07	4.70608404389006e-05
Applied Materials	9.01453283010842e-05	1.00875270008722e-05
Intel Corporation	0.00772339118521691	0.0185754413568268
Wells Fargo & Company	0.215669115079724	0.0306891722754909
Transocean LTD.	6.61834346503325e-05	2.52826362531821e-07
AT&T	0.00655344097245255	3.30204923710662e-07
Astoria Financial Corporation	4.98826639225408e-07	2.1236817134366e-08
Comcast Corporation	0.0167396966575649	3.1164339311266e-06
Brocade Communications Systems	0.0251767952214427	0.000997969872005882
KB-Home	0.0011439991546206	4.52547673286844e-05
Sirius XM Holdings	0.00019811644126806	0.0559368346097263
Verizon Communications	8.70437188730673e-07	9.51400787225313e-06
Marathon Oil Corporation	0.208204470825483	2.27474212206303e-07
Citigroup	0.160884642544214	0.157445357074501
Encana Corporation	0.120867363896224	3.46977568401131e-06
The Kroger Co.	8.60052427632613e-06	0.0234182850734007
Valbiant Pharmaceuticals International	6.62468444756966e-07	0.0463258040343431

FWA quote ottime: scenario 2		
	Configurazione 1	Configurazione 2
Rite Aid Corporation	0.00766519415494188	0.000466086381759426
Bank of America	1.06251032445096e-08	3.43400633250017e-06
Blackberry	6.5257365993003e-06	0.00964192983813947
Oracle Corporation	0.000650599316404956	0.0773688066504564
E.I. Du Pont de Nemours and Company	0.0475223546932314	1.71253252490513e-05
Ambev	0.000367204510663091	0.0122226122753237
Advanced Micro Devices	9.50853441100836e-07	0.0527509339548592
Micron Technology	2.53062926858642e-05	0.000393033033089289
General Eletric Company	1.16048512959629e-09	5.78958657459711e-05
Ford Motor Company	0.00535355935906336	6.89769778616259e-05
Vale	0.0575739122556817	0.0131112402393783
Apple	0.00055603077971897	0.000129491263139888
Chesapeake Energy Corporation	0.144653823138131	0.0439525142689016
Entreprise Products Partners L.P.	2.05272093504388e-07	5.34054779530996e-07
Applied Materials	9.01453283010842e-05	2.69728871114628e-07
Intel Corporation	0.00772339118521691	0.0631848378624128
Wells Fargo & Company	0.215669115079724	0.00986090039635979
Transocean LTD.	6.61834346503325e-05	0.163650836599047
AT&T	0.00655344097245255	0.0339904860692325
Astoria Financial Corporation	4.98826639225408e-07	0.0507654110928827
Comcast Corporation	0.0167396966575649	0.00198167506663305
Brocade Communications Systems	0.0251767952214427	3.95817364779067e-05
KB-Home	0.0011439991546206	0.000110461077541487
Sirius XM Holdings	0.00019811644126806	0.204641346574122
Verizon Communications	8.70437188730673e-07	0.000418048807939645
Marathon Oil Corporation	0.208204470825483	0.0148112885615499
Citigroup	0.160884642544214	0.00147876999009243
Encana Corporation	0.120867363896224	0.160219854969204
The Kroger Co.	8.60052427632613e-06	0.0125844606166629
Valbant Pharmaceuticals International	7.45700657956714e-07	2.36549804617812e-05

FWA quote ottime: scenario 3		
	Configurazione 1	Configurazione 2
Rite Aid Corporation	0.000163413789442845	0.00893085141085281
Bank of America	1.57678771735626e-07	0.000303857464432018
Blackberry	1.72693010941046e-06	0.0273560536152904
Oracle Corporation	2.2220585877362e-06	0.0296879868749788

E.I. Du Pont de Nemours and Company	4.23694552073519e-05	0.0934125640060023
Ambev	0.0500805987389131	0.00713812666743918
Advanced Micro Devices	4.29414962498335e-05	0.0588513483275833
Micron Technology	8.03921631569514e-06	0.0881102130591345
General Eletric Company	0.0387331605854327	0.000492517477116324
Ford Motor Company	0.0175589688288347	0.0265236956770259
Vale	0.0628565255182061	0.00092851933804666
Apple	2.21144151653925e-06	0.00258964285929544
Chesapeake Energy Corporation	0.152147486996573	0.158800626361538
Entreprise Products Partners L.P.	9.41847293926203e-06	0.000100058414573409
Applied Materials	0.21161388119298	0.00470372162673202
Intel Corporation	1.47955252036578e-06	0.00120394090271837
Wells Fargo & Company	0.108507502450907	0.000526936900705546
Transocean LTD.	0.142822071265393	0.000378389401204379
AT&T	2.98019744128433e-06	0.0366560608135029
Astoria Financial Corporation	0.00017994695605972	0.00130887857709213
Comcast Corporation	2.23492782046909e-05	0.00785404834970306
Brocade Communications Systems	0.0151067484783948	0.0399408291364979
KB-Home	0.000165163111824703	0.000274675804384692
Sirius XM Holdings	0.146994912460806	0.155313865851235
Verizon Communications	4.27115616311585e-08	0.146326122937244
Marathon Oil Corporation	7.76968076198682e-08	0.000388750168527374
Citigroup	0.000269596310777554	0.0995526213037573
Encana Corporation	0.000520351253370811	0.00272618368951941
The Kroger Co.	0.00308520397882558	0.00902677468635664
Valbant Pharmaceuticals International	4.51658416004795e-08	0.00000145416657445

PSO quote ottime: scenario 1		
Rite Aid Corporation	4.52195288179155e-40	
Bank of America	0.0339551291511783	
Blackberry	4.01466345400604e-48	
Oracle Corporation	0.0677547296063152	
E.I. Du Pont de Nemours and Company	0.0306754122719572	
Ambev	-1.82427508490829e-45	
Advanced Micro Devices	0.0333349660006544	
Micron Technology	0.0436158349238563	
General Eletric Company	3.82872748775718e-44	
Ford Motor Company	0.0425546701986977	
Vale	-4.3500450392131e-46	

Apple	0.0794190875423979
Chesapeake Energy Corporation	1.94595819382853e-44
Entreprise Products Partners L.P.	0.0314569287895941
Applied Materials	0.0313097764806718
Intel Corporation	0.0340932001670632
Wells Fargo & Company	0.0432516376927486
Transocean LTD.	2.22063697198064e-44
AT&T	0.0313817881423391
Astoria Financial Corporation	0.0537258969569718
Comcast Corporation	0.157265737176994
Brocade Communications Systems	0.0355376564113211
KB-Home	0.066774535388487
Sirius XM Holdings	0.0313285687070911
Verizon Communications	-2.03381940224715e-37
Marathon Oil Corporation	0.0825041314919899
Citigroup	-8.49990872874479e-48
Encana Corporation	0.039131174682863
The Kroger Co.	-9.79669462387843e-43
Valbant Pharmaceuticals International	1.7564841247555e-44

PSO quote ottime: scenario 2	
Rite Aid Corporation	4.48661199802174e-49
Bank of America	3.98295426305557e-48
Blackberry	0.16045769313396
Oracle Corporation	-5.52360991316707e-48
E.I. Du Pont de Nemours and Company	0.054980208506973
Ambev	0.089739093683543
Advanced Micro Devices	-7.22935505274246e-49
Micron Technology	6.84601648427527e-46
General Electric Company	0.0694158313204554
Ford Motor Company	1.51251614403952e-47
Vale	0.0316041682400167
Apple	0.198944393254703
Chesapeake Energy Corporation	0.0518627963661824
Entreprise Products Partners L.P.	0.0865229244503826
Applied Materials	6.10930752357455e-05
Intel Corporation	0.0363657633234178
Wells Fargo & Company	0.030387183371472
Transocean LTD.	-2.30173726506251e-47
AT&T	0.0300010208634992
Astoria Financial Corporation	0.0307823462534545
Comcast Corporation	3.98315113498059e-47

Brocade Communications Systems	-8.70167786944293e-49
KB-Home	-2.31521401649984e-49
Sirius XM Holdings	0.0302050255150677
Verizon Communications	0.0347417716046803
Marathon Oil Corporation	0.0299934166837404
Citigroup	0.0339160496903662
Encana Corporation	-2.60343498169505e-48
The Kroger Co.	-2.86033063888779e-08
Valbant Pharmaceuticals International	3.487441326654752e-47

PSO quote ottime: scenario 3	
Rite Aid Corporation	1.27722500741328e-42
Bank of America	0.0322425928109845
Blackberry	7.89141599277566e-43
Oracle Corporation	0.0408420599585399
E.I. Du Pont de Nemours and Company	-2.25534304212043e-44
Ambev	8.0362032012632e-43
Advanced Micro Devices	-6.3153256783625e-44
Micron Technology	0.0368209958836775
General Electric Company	0.0887753127462321
Ford Motor Company	0.0337088540038336
Vale	8.79313668805059e-41
Apple	0.203995369733784
Chesapeake Energy Corporation	0.152863059720351
Entreprise Products Partners L.P.	0.0391981685061668
Applied Materials	3.7287896898649e-43
Intel Corporation	0.0373972811613696
Wells Fargo & Company	-3.42807712204265e-45
Transocean LTD.	1.6212782421908e-42
AT&T	-7.04511739990343e-35
Astoria Financial Corporation	0.029978015787436
Comcast Corporation	9.20357996609334e-45
Brocade Communications Systems	0.0415258666274366
KB-Home	-7.97443056719427e-46
Sirius XM Holdings	0.0325772742538311
Verizon Communications	0.0359310479672838
Marathon Oil Corporation	0.0368520441402936
Citigroup	0.0419452083414598
Encana Corporation	0.0855251325143478
The Kroger Co.	2.47408303812147e-42

Valbant Pharmaceuticals International	4.824574512545001e-45
--	-----------------------

## Appendice B

Codice Matlab: 'Frontiera efficiente N>2'.

```
%FRONTIERA EFFICIENTE N>2
```

```
randn('state',2) %questo comando resetta il generatore, quindi il campionamento può essere replicato più volte ottenendo lo stesso risultato
```

```
%matrice di varianza-covarianza
```

```
dates = [today:today+29]'; %30 osservazioni per ogni titolo
```

```
data = [randn(1,30); randn(1,30) ;
```

```
randn(1,30);randn(1,30);randn(1,30);randn(1,30);randn(1,30);randn(1,30);randn(1,30);randn(1,30)]; %10 titoli e 30 osservazioni ognuno; il comando randn campiona la distribuzione normale standardizzata
```

```
f = fints(dates, data);
```

```
C = cov(f); %matrice varianza-covarianza
```

```
C_inv=inv(C); %inversa della matrice C
```

```
e=[1 1 1 1 1 1 1 1 1 1]'; %vettore di 1
```

```
mu=[mean(data)]'; %media rendimenti per ogni titolo
```

```
pigreco=-1:0.01:1; %vettore dei rendimenti desiderati
```

```
%scalari alfa,beta,epsilon nella formula del peso ottimo
```

```
alfa=mu'*C_inv*mu;
```

```
beta=e'*C_inv*mu;
```

```
epsilon=e'*C_inv*e;
```

```
%matrice pesi ottimi per diversi pigreco
```

```
x_optimum=((epsilon*C_inv*mu-beta*C_inv*e)*pigreco+(alfa*C_inv*e-
```

```
beta*C_inv*mu))/(alfa*epsilon-beta^2);
```

```
%vettore deviazioni standard del portafoglio per diversi pigreco
```

```
stdev_p=((epsilon*pigreco.^2-2*beta*pigreco+alfa)/(alfa*epsilon-beta^2)).^(1/2);
```

```
%grafico
```

```
plot(stdev_p,pigreco)
```

```
%vertice parabola
```

```
stdev_p_v=(1/epsilon)^(1/2)
```

```
r_p_v=(beta/epsilon)
```

```

% vertice evidenziato nel grafico
hold on
plot(stdev_p_v,r_p_v,'o')
hold off

```

Codice Matlab: 'Funzione utilità quadratica'.

### %FUNZIONE UTILITA' QUADRATICA

```

a=0.1; %coefficiente avversione al rischio
x=0:0.001:(1/a); %dominio dove la funzione è strettamente monotona positiva
y=x-((a/2)*(x.^2)); %funzione di utilità
plot(x,y) %grafico
grid on
set(gca,'xticklabel',[]) %elimina valori dell'asse x
set(gca,'yticklabel',[]) %elimina valori dell'asse y

```

```

hold on
x=(1/a):0.001:20; %dominio dove la funzione è strettamente monotona decrescente
y=x-((a/2)*(x.^2)); %funzione di utilità
plot(x,y)
hold off

```

Codice Matlab: 'Curve d'indifferenza'.

### %CURVE D'INDIFFERENZA

```

a=0.1; %coefficiente di avversione al rischio
mu=0:0.001:(1/a); %vettore rendimenti attesi
U=1; %primo livello di utilità

```

```

%Curva d'indifferenza per U=1
sigma_2=(-2/a*U+2*mu/a-mu.^2); %calcolo vettore di varianza
sigma=sqrt(sigma_2); %calcolo deviazione standard
plot(sigma,mu) %grafico
grid on
set(gca,'xticklabel',[]) %elimina valori dell'asse x
set(gca,'yticklabel',[]) %elimina valori dell'asse y

```

```

%Curva d'indifferenza per U=2

```

```

hold on
U=2;
sigma_2=(-2/a*U+2*mu/a-mu.^2);
sigma=sqrt(sigma_2);
plot(sigma,mu)
hold off

```

%Curva d'indifferenza per U=3

```

hold on
U=3;
sigma_2=(-2/a*U+2*mu/a-mu.^2);
sigma=sqrt(sigma_2);
plot(sigma,mu)
hold off

```

%Curva d'indifferenza per U=4

```

hold on
U=4;
sigma_2=(-2/a*U+2*mu/a-mu.^2);
sigma=sqrt(sigma_2);
plot(sigma,mu)
hold off

```

Codice Matlab: PSO

```

clc
clear
format long g
% Caricamento serie storiche e calcolo rendimenti logaritmici
[prezzi,testo] = xlsread('DATI_3_ecx.xlsx');
[T,numvar] = size(prezzi);
rend = log(prezzi(2:end,:)./prezzi(1:end-1,:));

%suddivisione tra in-sample e out of sample
oos=126; %numero rendiemnti +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); %rendiemnti medi out of sample

%dati di input
Rp=mean(rm_is); %rendimento obiettivo giornaliero
Kd=5; %numero minimo di titoli detenibili
Ku=30; %numero massimo di titoli detenibili
d=ones(1,numvar)*0.03; %percentual d'investimento minima in ciasun titolo
u=ones(1,numvar)*0.2; %percentuale d'investimento massima in ciasun titolo

```

```

%Parametri
P=50; %numero particelle
c1=1.49618;
c2=1.49618;

a=1;
b=2;
niter=1000;
epsilon=0.0000001;
iw=0.7298;%peso d'inerzia

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

TT=(T-oos);

% vincoli
%R_is=zeros(P,1); %rendiemnto di portafoglio
%R_oos=zeros(P,1);
rho=zeros(P,1);
rm_rho=zeros(P,1);
R_is=zeros(P,1); %rendiemnto di portafoglio
R_oos=zeros(P,1);
vinc_1=zeros(P,1); %vincolo di bilancio
vinc_2=zeros(P,1); %vincolo sul rendimento minimo atteso
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 e 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

```

```

%for p=1:P
% for i=1:numvar
% x1(p,i)=x(p,i)*z(p,i);
% end
%end

```

```

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

```

```

%R_is(p)=x(p,:)*rm_is';
%R_oos(p)=x(p,:)*rm_oos';

```

```

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

```

```

rm_rho(p)=x1(p,:)*rm_is';
rho(p)=((a/TT)*(sum(max(0,prodotto_somma_is(p,:))))+(1-a)*(((1/TT)*(sum((max(0,-
prodotto_somma_is(p,:)).^b)))^(1/b))-rm_rho(p));

```

```

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'integrità
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+vinc_8));
```

```
%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
```

```
%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

plot(converg)

quote_ottime=gx(1,1:numvar)'
best_fitness=gx(1,numvar+1)

rho_is=rho(posizione)
rho_oos=(a/(T-TT-1))*(sum(max(0,differenza_oos*quote_ottime)))+(1-a)*(1/(T-TT-1))*(sum(max(0,-(differenza_oos*quote_ottime)).^b)^(1/b))-(quote_ottime'*rm_oos')
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)
vinc_non_negativita=vinc_8(posizione)

```

Codice Matlab: FWA

```
%FWA
```

```
clear
```

```
clc
```

```
format long g
```

```

% Caricamento serie storiche e calcolo rendimenti logaritmici
[prezzi,testo] = xlsread('DATI_1_exc.xlsx');
[T,numvar] = size(prezzi);
rend = log(prezzi(2:end,:)./prezzi(1:end-1,:));

%suddivisione tra in-sample e out of sample
oos=126; %numero rendiemnti +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); %rendiemnti medi out of sample

TT=T-oos;

%Dati di input
Rp=mean(rm_is); %rendimento obiettivo giornaliero
Kd=7; %numero minimo di titoli detenibili
Ku=30; %numero massimo di titoli detenibili
d=ones(1,numvar)*0.03; %percentual d'investimento minima in ciasun titolo
u=ones(1,numvar)*0.2; %percentuale d'investimento massima in ciasun titolo

a=0.5;
b=2;

%parametri FWA
N=5; %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 calocolo del numero di
scintille sia zero
m_gauss=1; %numero di scintille prodotto dall'esplosione Gaussiana
niter=1000; %numero d'iterazioni
x_max=100; %limite superiore dello spazio
x_min=-100; %limite inferiore dello spazio
x_max_z=100;
x_min_z=-100;

%vincoli, vettori di appoggio per la funzione obiettivo
%R_is=zeros(N,1); %rendiemnto di portafoglio
%R_oos=zeros(N,1);
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 ciasun titolo i per ciasuna 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);

```

```

%f=ones(N,1)*1.0e+255; %funzione fi fitness
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
rm_rho(n)=x1(n,:)*rm_is';
rho(n)=((a/TT)*(sum(max(0,prodotto_somma_is(n,:))))+(1-a)*(((1/TT)*(sum((max(0,-
prodotto_somma_is(n,:)).^b)))^(1/b))-rm_rho(n));

```

```

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
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_3+vinc_4+vinc_5+vinc_6+vinc_7+v
inc_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 scintile 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
```

```

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
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';

    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_3_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)

```

**%RISULATI**

```

f(1)
quote_ottime=x1(1,:)
rho_oos=((a/(T-TT-1))*(sum(max(0,differenza_oos*quote_ottime))))+(1-a)*(((1/(T-TT-1))*(sum((max(0,-(differenza_oos*quote_ottime)).^b))))^(1/b))-(quote_ottime'*rm_oos'))
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_non_negativita=vinc_3_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)
```

## Bibliografia

- Eldwin, E. J., Gruber, M. J., Brown, S. J., Goetzmann, W. N.: *Teorie di portafoglio e analisi degli investimenti*, Apogeo, 2007.
- Tan, Y.: *Fireworks Algorithm. A novel Swarm Intelligence Optimization Method*, Springer, 2015
- Peirò, A.: *Skewness in financial returns*, Journal of Banking and Finance, 23 (6), 847-862, 1999.
- Scott, R.C., Howarth, P.A.: *On the direction of preference for moments of higher order than the variance*. Journal of Finance, 35 (4), 915-919, 1980.
- Wipperfurth, R.F.: *Utility implications of portfolio selection and performance appraisal models*, Journal of Financial and Quantitative Analysis, 6 (3), 913-924, 1971.
- Tan, Y., Zhu, Y.: *Fireworks Algorithm for Optimization*, in Advances in Swarm Intelligence, Springer, 2010.
- Radovan, A.: *La PSO per la selezione di portafogli basati sul Partitioned VaR e su vincoli misti-interi*, Università Ca' Foscari, Venezia, 2014
- Blum, C., Roli, A.: *Metaheuristics in Combinatorial Optimization: Overview and Conceptual Comparison*, ACM Computing Survey, 2003.
- 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, 611-624, 2013.
- Sorensen, K., Sevaux, M., Glover, F.: *A history of Metaheuristics*, in Mart, R., Pardalos, P., Resende, M., Handbook of Heuristics, Springer 2016.
- Bacanin, N., Tuba, M. *Fireworks Algorithm Applied to Constrained Portfolio Optimization Problem*, in IEEE Congress on Evolutionary Computation (CEC), 2015.
- Peram, T., Veeramachaneni, K., & Mohan, C.: *Fitness-distance ratio based particle swarm optimization*, in IEEE swarm intelligence symposium (SIS), 2003.

Poli, R., Kennedy, J., Blackwell, T.: *Particle Swarm Optimization. An overview*, in *Swarm Intelligence*, 1 (1), 33–57, 2007.

Chen, Z., Wang, Y.: *Two-sided coherent risk measures and their application in realistic portfolio optimization*, *Journal of Banking and Finance* 32(12), pp. 2667-2673, 2008.

Conine, T.E. Jr, Tamarkin, M.J.: *On diversification given asymmetry in returns*, *Journal of Finance* 36 (5), 1143-1155, 1981.

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, 611-624, 2013.

Begotto, M. *Approccio evolutivo e analisi delle serie storiche*, Tesi di Laurea, Università Ca' Foscari Venezia, 2014.

Bruno, G.: *Algoritmi euristici, costruttivi e migliorativi*, scaricato dal sito <http://www.federica.unina.it> in data 5 settembre 2017.

Marchi, G: *Principi di gestione di portafoglio e teoria di Markowitz*, scaricato dal sito <http://jobtrading.it> in data 20 agosto 2017.

Saltari, E.: *Il criterio media-varianza e il modello CAPM*, scaricato dal sito [http://www.webalice.it/esaltari/Didattica/Economia%20Finanziaria%20I/capm\\_fincap05.pdf](http://www.webalice.it/esaltari/Didattica/Economia%20Finanziaria%20I/capm_fincap05.pdf)

Codogno, M.:  *$P! = NP$  (o no?)*, scaricato dal sito <http://www.ilpost.it/mauriziocodogno> in data 20 agosto 2017.

Corazza, M.: corso di Metodi per la gestione di portafogli, Università Ca' Foscari, Venezia, 2016.

Saltari, E., *Il modello media-varianza con  $N$  titoli rischiosi. Una derivazione formale*, scaricato dal sito <http://www.webalice.it/esaltari/Didattica/Economia%20Finanziaria%20I/ntitoli.pdf> in data 20 luglio 2017.

Fumagalli, T.: *La teoria della selezione di portafoglio di Markowitz*, scaricato dal sito <http://www00.unibg.it> in data 2 agosto 2017.