Università
Ca'Foscari
Venezia

Corso di Laurea Magistrale

in Economia e Finanza
(ordinamento ex D.M. 270/2004)

Tesi di Laurea

PSO e BFO per la gestione di hedge funds: il caso DeltaHedge

Relatore
Ch. Prof. Marco Corazza

Laureando
Michael Rizzo
Matricola 849687

Anno Accademico
2017 / 2018

A Michele,

senza dubbio.

Indice

Introduzione ... 1

1 MODERN PORTFOLIO THEORY ... 3

1.1 Aspetti introduttivi ... 3

1.2 Principio di scelta basato sul criterio media-varianza ... 5

1.3 Dalla teoria dell'utilità attesa al portafoglio ottimo ... 14

1.4 Critiche al modello di selezione alla Markowitz ... 20

2 DALLE MISURE DI RISCHIO PREMODERNE ALLE MISURE DI RISCHIO COERENTI ... 29

2.1 Il Value at Risk ... 29

2.1.1 Metodi di stima del Value at Risk ... 33

2.1.2 Limiti del Value at Risk ... 35

2.2 Gli assiomi di coerenza ... 38

2.3 Alcune misure di rischio coerenti ... 41

2.3.1 Tail Conditional Expectation e Worst Conditional Expectation ... 41

2.3.2 Expected shortfall e Conditional Value at Risk ... 42

2.4 Misure di rischio two-sided ... 45

3 PARTICLE SWARM OPTIMIZATION E BACTERIAL FORAGING OPTIMIZATION ... 49

3.1 Euristiche ed algoritmi esatti ... 49

3.2 Metaeuristiche ... 51

3.2.1 Gli algoritmi evolutivi ... 54

3.2.2 Swarm Intelligence ... 55

3.3 Particle Swarm Optimization ... 57

3.3.1 L'algoritmo originale ... 60

3.3.2 Varianti della PSO ... 61

3.3.3 Topologia della popolazione 64

3.4 Bacterial Foraging Optimization 67

3.4.1 BFO per la risoluzione di problemi di ottimizzazione di portafoglio 72

4 IL CASO DELTAHEDGE 77

4.1 La società 77

4.2 I futures 79

4.3 Presentazione del modello di selezione 85

4.4 Riformulazione del modello di selezione 91

5 DETERMINAZIONE DELL'ALLOCAZIONE FINALE ALL'INTERNO DI STYX 95

5.1 Informazioni iniziali 95

5.2 Impostazione dei parametri 100

5.3 Applicazione delle metaeuristiche 103

5.4 Confronto ed analisi 106

5.5 Determinazione dell'allocazione finale 113

Conclusioni 117

Appendice A 119

Appendice B 120

Appendice C 121

Appendice D 137

Bibliografia 139

Introduzione

L'obiettivo di questo lavoro è quello di risolvere un reale problema di ottimizzazione che ha dovuto affrontare la società DeltaHedge al momento dell'avvio del suo *managed futures program* Styx. Quest'ultimo investe in nove strategie di trading contrarian intraday, le quali differiscono sulla base delle tipologie di contratti futures sui quali operano. In particolare, ogni singola strategia effettua le proprie operazioni di acquisto e vendita su uno dei seguenti futures: 6A, 6B, 6C, 6E, 6J, 6M, 6N, 6S, DX.[1] In altre parole, dato un capitale iniziale di 2 milioni di dollari, si vuole determinare l'allocazione ottima, da un punto di vista rischio-rendimento, all'interno di queste nove strategie che compongono Styx. Inoltre, l'allocazione ottenuta non deve semplicemente indicare le percentuali di capitale da attribuire ad ogni sistema di trading, ma deve specificare il numero di contratti futures mediante il quale ogni singola strategia deve operare. Cambiando la prospettiva, questo problema di minimizzazione può essere visto come una specifica tipologia di *asset allocation*, dove l'insieme delle *asset class*, sulle quali è possibile investire, sono differenti strategie di trading. Il modello di selezione, risultante dalla modellizzazione delle richieste avanzate da DeltaHedge, presenta una misura di rischio ed un sistema di vincoli che lo differenziano dal classico modello proposto da Markowitz, il padre della *Modern Portfolio Theory*. Più in dettaglio, per quanto riguarda la funzione obiettivo, la scelta è ricaduta su una misura di rischio two-sided, la quale, oltre a considerare l'intera distribuzione dei rendimenti, è una misura che rispetta i quattro assiomi di coerenza: l'invarianza transizionale, la subadditività, la positiva omogeneità e la monotonicità. Il rispetto di queste quattro proprietà assicura una forma di coincidenza tra l'incertezza percepita dal singolo investitore razionale ed il numero che la riassume mediante un'apposita funzione. Come si vedrà nel seguito, l'impiego di questa misura di rischio consente di superare molti limiti connessi con l'utilizzo della varianza come funzione obiettivo da minimizzare. Passando ora ai vincoli applicati al modello, questi permettono di ottenere un'allocazione che sia realmente adottabile da parte di DeltaHedge in Styx. I vincoli

[1] In Appendice A la legenda delle sigle dei futures.

impostati sono i seguenti: vincolo sul profitto medio minimo, vincolo sulla leva finanziaria, vincolo di cardinalità, vincolo sulla non negatività del numero di contratti, vincolo d'integralità e il vincolo sulle percentuali di investimento. Se da un lato tali vincoli rendono il modello di selezione "reale" da un punto di vista applicativo, dall'altro lato lo rendono piuttosto complesso, al punto che l'utilizzo di algoritmi esatti per la sua risoluzione è una strada non percorribile.

Per la risoluzione di questo problema di minimizzazione vincolata si è deciso di fare ricorso a due metaeuristiche, ovvero degli algoritmi che, oltre ad avvicinarsi alla soluzione ottima, cercano di esplorare in modo più dettagliato quelle aree che risultano più interessanti all'interno dello spazio delle soluzioni. La prima è ispirata al movimento degli stormi di uccelli in situazioni di "emergenza" o nella ricerca del cibo, ossia la Particle Swarm Optimization (PSO). La seconda, definita Bacterial Foraging Optimization (BFO), prende ispirazione dal comportamento dei batteri Escherichia coli. Dato che entrambe le metaeuristiche appena definite sono applicabili solamente a problemi di ottimizzazione libera, si ricorrerà al metodo delle penalità esatte per riformulare il problema iniziale in forma non vincolata.

Nell'ultimo capitolo, si procederà con la determinazione dell'allocazione "ottima" per il programma Styx seguendo un processo che può essere riassunto in due passi: prima si verificherà se una metaeuristica offre soluzioni migliori rispetto all'altra osservando le performance generate nei differenti futuri virtuali; in seguito si utilizzerà la metaeuristica "migliore" con il fine di trovare l'allocazione ottima finale per Styx,

L'applicabilità di tale modello supera i confini delimitati dallo specifico caso oggetto di questo lavoro, e può essere adattato a qualsiasi situazione in cui una determinata società deve decidere come distribuire un certo capitale all'interno di un fondo che investe su un paniere di futures attraverso differenti strategie di trading.

CAPITOLO 1

MODERN PORTFOLIO THEORY

In questo primo capitolo si intende presentare la teoria di selezione del portafoglio proposta da Markowitz. Una volta elencate l'insieme di ipotesi su cui si fonda quest'ultima, si illustrerà la procedura da seguire per la costruzione della frontiera efficiente di un portafoglio composto da N titoli rischiosi. Successivamente verrà introdotta la teoria dell'utilità attesa, la quale consente di identificare il portafoglio ottimo tra tutti quelli che formano la frontiera efficiente. Nel paragrafo conclusivo si vedranno i limiti che caratterizzano questo metodo di selezione.

1.1 Aspetti introduttivi

"Investire" è la fase conclusiva della teoria microeconomica del consumatore: un individuo che al tempo *t* consuma solo una parte del suo capitale, massimizzando la sua utilità, deve "trasportare" il capitale rimanente all'istante successivo, investendolo in uno o più *asset.*

In particolare, la scelta di ripartire la ricchezza disponibile tra due o più attività finanziarie porta alla formazione di un portafoglio ossia, da un punto di vista quantitativo, un vettore di dimensioni *NX1*, dove ogni elemento definisce la quota di capitale allocata a uno dei possibili *N asset* presenti sul mercato:

$$x' = (x_1, x_2, \dots, x_N).$$

La composizione di quest'ultimo deve essere determinata mediante un accurato processo di selezione, poiché si possono ottenere performance out-of-sample molto differenti variando, anche solo leggermente, la composizione del portafoglio su cui si è investito.

Il contributo più importante alla teoria di selezione del portafoglio è stato fornito da Harry Max Markowitz quando nel 1952, all'età di soli 25 anni, pubblicò il paper

"Portfolio Selection", creando le fondamenta di quella che oggi viene definita come *Modern Portfolio Theory.* La rilevanza del modello di selezione sviluppato da Markowitz è amplificata dal fatto che tutt'ora questo viene utilizzato da un numero considerevole di gestori di fondi sia per determinare l'allocazione efficiente, sia come benchmark per confrontare modelli alternativi.

Prima di procedere ad una descrizione più dettagliata di tale modello è necessario definire il sistema di ipotesi su cui questo si fonda, in quanto essenziali per circoscrivere l'ambiente di lavoro:

- i mercati finanziari sono senza frizioni (frictionless markets assumption), cioè non ci sono né tasse né costi di transizione e tutti gli asset sono perfettamente divisibili;
- gli investitori sono *price-taker* (price-taker assumption), ossia nessun individuo, attraverso il suo comportamento, può influenzare la distribuzione di probabilità dei rendimenti degli *asset* disponibili;
- non ci sono vincoli istituzionali in quanto lo *short-selling*[2] è ammesso (no-istitutional restrictions assumption);
- gli agenti economici conoscono la distribuzione di probabilità dei rendimenti degli asset;
- l'economia è articolata in due soli istanti temporali (*t* e *t+1*), dove nel primo un individuo sceglie in quali *asset* investire la propria ricchezza, mentre nel secondo verifica i propri guadagni o perdite conseguite;
- principio di *non sazietà* degli investitori poiché, a parità di condizioni, preferiranno sempre un livello di ricchezza maggiore rispetto a uno minore;
- gli investitori sono avversi al rischio, ovvero dovendo scegliere tra due titoli con lo stesso rendimento atteso, sceglieranno sempre quello con il rischio minore.

È definito *razionale* un investitore che rispetta entrambi gli assiomi di *non sazietà* e

[2] Lo *short-selling*, o vendita allo scoperto, consiste nel prendere a prestito un titolo da un investitore, spesso pagando una certa percentuale di commissioni, con lo scopo di venderlo ad un altro soggetto: in altre parole, si vende un titolo che in realtà ancora non si possiede. Nel momento in cui il venditore allo scoperto dovrà restituire il titolo al suo proprietario, sarà obbligato ad acquistare tale attività finanziaria sul mercato. Tutta questa operazione genererà un guadagno per il venditore allo scoperto solamente se il prezzo del titolo preso a prestito dovesse scendere.

di avversione al rischio.
È necessario evidenziare che le ipotesi sono importanti poiché permettono di capire quando è possibile applicare un dato modello ed entro quali limiti esso è valido. Tuttavia non tutte le ipotesi definite sono valide nel mondo reale, è proprio per questo che si parla di "modelli".

1.2 Principio di scelta basato sul criterio media-varianza

Il modello di selezione proposto da Markowitz può essere definito come un processo composto da tre passi che porta alla costruzione di un portafoglio ottimo.
La prima fase consiste nel definire gli strumenti che hanno l'obbiettivo di misurare la redditività e il rischio di un determinato portafoglio. Il problema è che la maggior parte degli *asset* finanziari hanno un rendimento aleatorio, cioè non è possibile conoscere al tempo t quanto effettivamente verrà realizzato (in termini di rendimento) da un determinato investimento in un periodo futuro ($t+1$). Di conseguenza, si può affermare che il rendimento di un titolo è una variabile casuale con una certa distribuzione di probabilità.
Markowitz, con l'obbiettivo di catturare l'incertezza associata a un determinato investimento, propose una coppia di indici statistici legati alla variabile casuale rendimento: il valore atteso come misura della redditività futura e la varianza come misura di rischio di tale redditività. Quindi questo modello ipotizza che tutte le peculiarità di un titolo possano essere sintetizzate dai primi due momenti della sua distribuzione di probabilità dei rendimenti.
In termini generali, considerando inizialmente un solo titolo, il valore atteso di una variabile casuale discreta X è dato semplicemente dalla formula:

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

dove:

- x_i, con $i = 1,2, \dots, M$, è la *i-esima* realizzazione di X;
- p_i con $0 \leq p \leq 1$ per ogni i e $\sum p_i = 1$, è la probabilità che si verifichi l'evento x_i.

Inoltre, nel caso di variabile casuale continua caratterizzata da una funzione di densità $f_x(.)$, il valore atteso di *X* è dato dall'integrale:

$$E(X) = \int_{-\infty}^{\infty} tf(t)dt\,;$$

in quest'ultimo caso *E(X)* potrebbe non esistere.

In termini generici, la media di una variabile casuale è un indice statistico di posizione; tuttavia, da un punto di vista finanziario, si considera tale indice come una misura della profittabilità di una scelta d'investimento.
Per quanto riguarda la varianza di una variabile discreta *X*, questa è data dalla somma ponderata del quadrato degli scostamenti delle singole realizzazioni rispetto alla media:

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

Nel caso di variabile casuale continua caratterizzata da una funzione di densità $f_x(.)$, la varianza di *X* è calcolata come:

$$Var(X) = \int_{-\infty}^{\infty} ((t - E(X))^2 f(t)dt\,;$$

da un punto di vista finanziario, la varianza dei rendimenti di un investimento è considerata come una misura di rischio, nonostante sia stata sviluppata come indice di variabilità. Come per il valore atteso, $Var(X)$ potrebbe non esistere.
A questo punto si può illustrare come determinare il rendimento atteso e la varianza di un portafoglio composto da più titoli; la specificazione di tali misure è indispensabile per passare alla fase successiva.
Partendo dal rendimento del portafoglio R_p, questo è dato dalla somma pesata dei singoli rendimenti aleatori dei titoli che lo compongono,

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

dove:

- x_i definisce il peso di ogni titolo presente nel portafoglio,
- R_i sono delle variabili casuali che rappresentano la redditività aleatoria dei vari titoli.

Applicando alcune proprietà del valor medio, si ottiene la media del rendimento del portafoglio come

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

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

dove r_i è il rendimento medio del titolo *i-esimo*. Esiste quindi una relazione lineare tra le percentuali di capitale investite nei titoli e la redditività del portafoglio.
Per quanto concerne la varianza dei rendimenti di un portafoglio, questa è più complessa da calcolare, poiché è pari alla somma pesata delle varianze dei rendimenti dei titoli che lo compongono, solamente nel momento in cui questi non hanno alcun legame tra di loro (linearmente indipendenti). Tuttavia, nei mercati finanziari, la maggior parte delle azioni presenta qualche livello di dipendenza (in particolare quelle appartenenti allo stesso settore), quindi è errata l'assunzione dell'ipotesi di indipendenza.
Per misurare la dipendenza nell'ambito della varianza si utilizza l'operatore covarianza, ossia uno strumento statistico in grado di catturare le relazioni lineari tra le singole coppie di titoli che compongono il portafoglio. Tale operatore si rende indispensabile per il calcolo della varianza del rendimento di un portafoglio:

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

Come si può vedere, la varianza è in parte determinata dalla varianza del rendimento dei singoli titoli e in parte dalle covarianze esistenti tra tutte le coppie di titoli. Inoltre, può essere espressa in termini matriciali come prodotto di tre matrici:

$$\sigma_P^2 = x'Vx,$$

dove x' e x sono rispettivamente il vettore riga e il vettore colonna dei relativi pesi, mentre V è la matrice varianza-covarianza[3]. Il risultato di tale prodotto matriciale non può essere negativo, in quanto la varianza è sempre maggiore o uguale a zero. La covarianza, all'opposto, varia da $-\infty$ a $+\infty$, perciò il suo utilizzo a fini comparativi può risultare complesso. Per ovviare a questo problema si utilizza il coefficiente di correlazione, che presenta il vantaggio di variare all'interno dell'intervallo -1 e +1:

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

quindi la varianza del rendimento del portafoglio può essere riformulata come

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

Una peculiarità della covarianza (o della correlazione), nonché un suo limite, è la capacità di catturare solo le dipendenze lineari, non considerando quelle relazioni tra titoli modellizzate da una funzione non lineare: ad esempio, ci sono variabili casuali che presentano una covarianza pari a zero nonostante una relazione tra di loro del tipo non lineare. Detto ciò, si può affermare che l'indipendenza implica covarianza (o correlazione) nulla, ma covarianza (o correlazione) nulla non implica indipendenza; pertanto quando i mercati sono dominati da dinamiche complesse, l'approccio media-varianza può funzionare peggio rispetto ad un mercato meno agitato.

Esposti gli strumenti attraverso i quali misurare l'incertezza associata a una data scelta di investimento, si può avanzare alla seconda fase del processo di selezione del portafoglio ottimo, ossia la definizione di un criterio efficiente che mediante tali strumenti sia in grado di determinare quali portafogli non sono dominati da altri secondo tale criterio. A questo fine, Markowitz propose il criterio media-varianza: date due variabili casuali X_1 e X_2, si può affermare che X_1 domina X_2 (o X_1 è preferito a X_2) in media-varianza se

[3] Matrice che presenta le varianze e le covarianze dei titoli rispettivamente lungo e al di fuori della diagonale.

- $E(X_1) \geq E(X_2)$ e
- $Var(X_1) \leq Var(X_2)$ e
- *almeno una di tali disuguaglianze* è *soddisfatta in senso stretto*.

A titolo esemplificativo, basti pensare a tre portafogli P_1, P_2 e P_3, dove il primo presenta un rendimento medio del 10% e deviazione standard del 3%, il secondo un rendimento medio del 5% e deviazione standard del 4%, mentre il terzo un rendimento medio del 4% e deviazione standard del 4%: come si può vedere dalla Figura 1, P_1 domina sia P_2 che P_3 in media-varianza poiché presenta una media superiore e un rischio inferiore rispetto al secondo e al terzo.

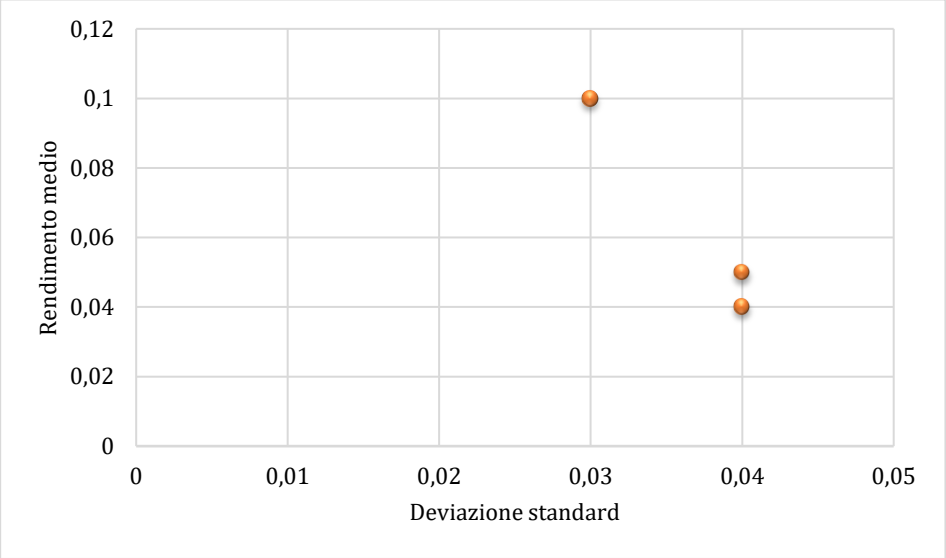


Figura 1: confronto in media-varianza

È evidente che tale criterio di scelta non è completo, difatti, in presenza di due portafogli entrambi efficienti, esso non è in grado di definire quale dei due sia preferibile all'altro. Di conseguenza, formano la cosiddetta *frontiera efficiente* tutti i portafogli ottimi, cioè quelli per cui fissato un certo livello di rendimento atteso minimizzano la varianza, oppure dato il livello di varianza massimizzano il rendimento atteso. Perciò tutti i portafogli che si trovano su tale frontiera saranno efficienti e saranno gli unici su cui ricadrà la scelta di un individuo razionale.
Markowitz fu il primo a proporre un modello di programmazione quadratica lineare per la formulazione del problema della costruzione della *frontiera efficiente*:

$$min_{x_1,\dots,x_n}\ \sigma_p^2$$

$$s.t. \begin{cases} E(r_p) = \pi \\ \sum_{i=1}^{n} x_i = 1 \end{cases}$$

o con forma matriciale: $$min_{x_1,\dots,x_n}\ x'Vx$$

$$s.t. \begin{cases} x'r = \pi \\ x'e = 1 \end{cases},$$

dove:

- π è il valore del rendimento atteso desiderato dell'individuo;
- x rappresenta il vettore di ordine N che contiene le percentuali di capitale investite in ogni titolo;
- V è la matrice varianza-covarianza;
- r è il vettore di ordine N che include i rendimenti attesi di tutti i titoli;
- e è un vettore unitario di ordine N.

Più dettagliatamente, questo modello è un problema di minimizzazione vincolata, poiché minimizza il rischio nel rispetto di due vincoli lineari: il primo definisce il rendimento atteso che si deve ottenere dal portafoglio selezionato, mentre il secondo esige che tutto il capitale disponibile sia investito (il cosiddetto vincolo di bilancio).

Prima di presentare l'algoritmo utilizzato per risolvere tale problema di ottimizzazione, è necessario definire le assunzioni che devono essere soddisfatte affinché esista una soluzione univoca, rappresentata dal vettore contenente i pesi ottimi attribuiti ad ogni titolo. La prima è che la matrice varianza-covarianza dev'essere non singolare[4], ossia tutte le righe e colonne della matrice devono essere linearmente indipendenti. In termini finanziari, questo significa che non devono esistere titoli con strutture di rischio esprimibili come combinazioni lineari delle strutture di rischio di altri titoli. La seconda condizione è che la matrice varianza-covarianza deve essere definita positiva, cioè tutte le varianze dei vari titoli devono essere strettamente maggiori di zero. Infine, l'ultima condizione è che ci siano almeno due titoli che presentano un rendimento atteso differente uno dall'altro, altrimenti il problema diventerebbe banale, perché tutta la ricchezza verrebbe attribuita al titolo con varianza inferiore.

[4] Una matrice non singolare è invertibile.

A questo punto si può definire il processo di soluzione del problema, che consiste nel differenziare la seguente Lagrangiana nel rispetto di ogni variabile ed ai moltiplicatori di Lagrange (λ_1 e λ_2), e ponendo le equazioni risultanti pari a zero:

$$L = x'Vx - \lambda_1(x'r - \pi) - \lambda_2(x'e - 1),$$

risolvendo il sistema delle condizioni di primo ordine

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

da cui dopo una serie di passaggi, si arriva alla seguente soluzione

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

dove

- x^* è il vettore contenente i pesi attribuiti ai singoli titoli che compongono il portafoglio ottimo selezionato,
- $\alpha = \mu' V^{-1}\mu$,
- $\beta = \mu' V^{-1}e = e' V^{-1}\mu$,
- $\gamma = e' V^{-1}e$.

Tale formula consente di determinare i portafogli efficienti per ogni valore del rendimento desiderato π e quindi di rappresentare la frontiera efficiente. Inoltre, riformulando i risultati precedenti e considerando un numero di titoli maggiore o uguale a due, l'espressione della frontiera può essere espressa in termini di varianza come

$$\sigma_P^2 = x^{*\prime} V x^* = \frac{\gamma\pi^2 - 2\beta\pi + \alpha}{\alpha\gamma - \beta^2},$$

o in termini di deviazione standard come

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

Date queste due formule, si può derivare la forma della frontiera sia nel piano media-varianza che nel piano media-deviazione standard: nel primo caso è una parabola con coordinate del vertice $\left(\sigma^2_{P,v} = \frac{1}{\gamma}, r_{P,v} = \frac{\beta}{\gamma}\right)$, mentre nel secondo caso è un'iperbole con coordinate del vertice $\left(\sigma_{P,v} = \left(\frac{1}{\gamma}\right)^{\frac{1}{2}}, r_{P,v} = \frac{\beta}{\gamma}\right)$.

La Figura 2 mostra entrambe le frontiere efficienti ottenute con Matlab, partendo dai prezzi giornalieri (190 osservazioni) di cinque titoli azionari scelti casualmente tra quelli presenti in Borsa Italiana[5]. In entrambi i casi, la frontiera efficiente è rappresentata dal lato superiore della linea continua azzurra che parte dal vertice (cerchio grigio), ossia da quel punto che definisce il portafoglio ottimo con minor rischio; tutti i portafogli che si trovano al di sotto di tale frontiera sono inefficienti.
I punti rossi figurano il rendimento atteso e la varianza (o deviazione standard) corrispondenti ai cinque titoli presi singolarmente.

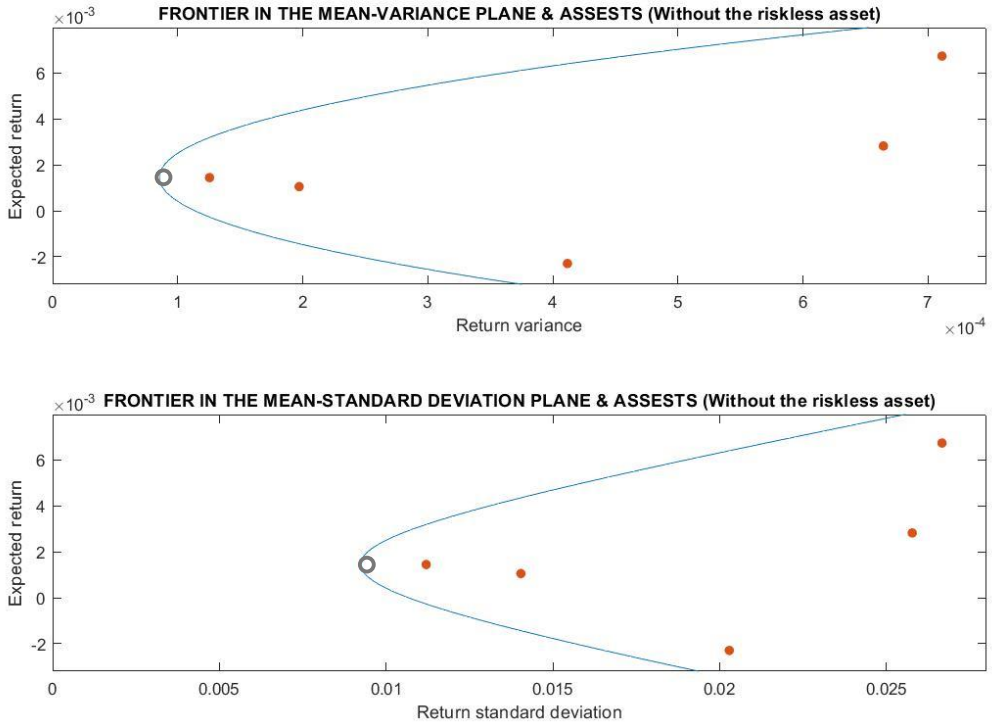


Figura 2: frontiera efficiente nel piano media-varianza e media-deviazione standard con N=5.

[5] Saipem, Campari, Carraro, Intesa SP e Fiat.

L'aspetto più interessante della *Modern Portfolio Theory* è sicuramente il concetto di diversificazione. Tale principio parte dall'idea che, investendo in due o più titoli che presentano covarianza inferiore a uno, può essere possibile contrarre il rischio complessivo del portafoglio ad un livello inferiore rispetto a quello del titolo meno rischioso. Quindi il semplice investire in più titoli, ossia diversificare, permette di ridurre il rischio di un portafoglio.

Per chiarire tale concetto si utilizza il portafoglio detto equally-weighted, ossia quel portafoglio che ripartisce le stesse percentuali di capitale in ogni titolo su cui investe. La varianza dei rendimenti di tale portafoglio può essere definita come:

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

Entrambe le sommatorie sono delle medie: la prima rappresenta la varianza media dei titoli che compongono il portafoglio, mentre la seconda indica la covarianza media degli stessi. Questa formula può essere riformulata come:

$$\sigma_P^2 = \frac{1}{N}\bar{\sigma}_i^2(N) + \frac{N-1}{N}\bar{\sigma}_{i,j}(N),$$

dove si può notare che quando *N* tende ad infinito, da un lato il primo prodotto tende a zero, poiché il rapporto $\frac{1}{N}$ tende a zero, dall'altro lato il secondo prodotto tende a $\bar{\sigma}_{i,j}(N)$, in quanto il rapporto $\frac{N-1}{N}$ tende ad uno.

Questo significa che investendo in un numero molto elevato di titoli, al limite infinito, la varianza del portafoglio è determinata quasi interamente dalla covarianza media, dato che la varianza media tende a scomparire. Esiste una relazione inversa tra il numero di titoli che compongono il portafoglio e il valore della varianza dei rendimenti dello stesso. La Figura 3 esemplifica quanto detto evidenziando, tuttavia, l'esistenza di una soglia di rischio non eliminabile detta *rischio idiosincratico*.

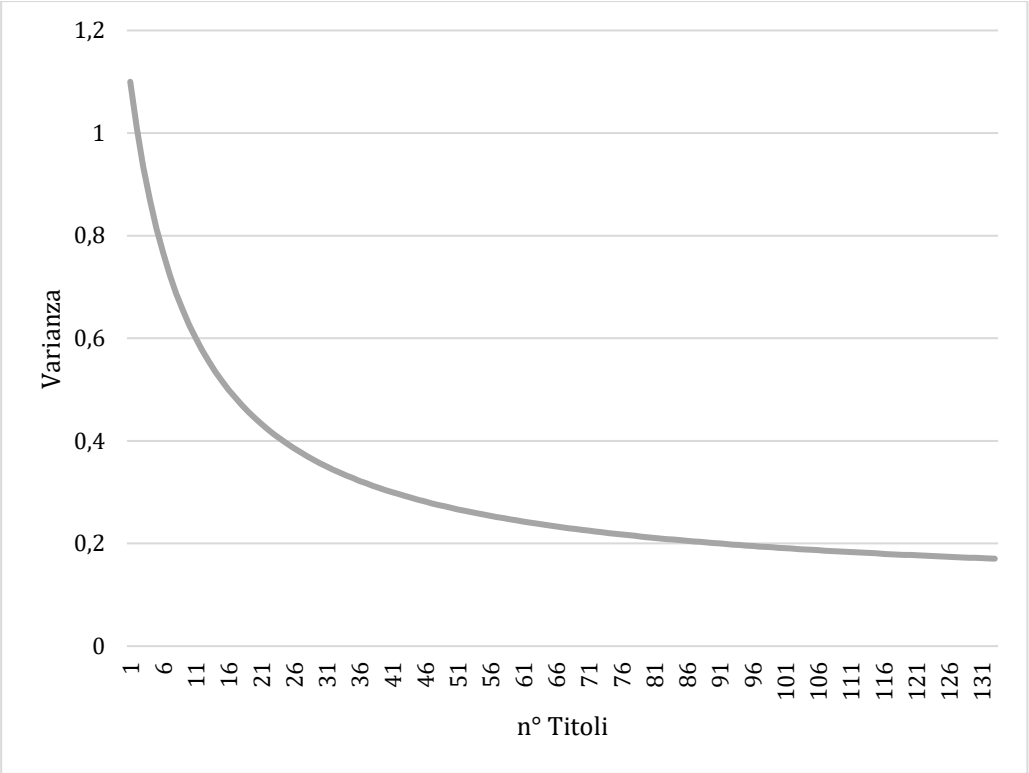


Figura 3: relazione tra varianza del portafoglio e il numero di titoli che lo compone.

Nella realtà non esistono infiniti titoli, e se anche questo fosse vero, ci sarebbe una quantità troppo elevata di costi da sostenere per il loro acquisto.

La terza e ultima fase del processo di selezione alla Markowitz consiste nel selezionare, tra tutti portafogli appartenenti alla frontiera efficiente, quello che meglio soddisfa il profilo di rischio-rendimento del singolo investitore. Per raggiungere tale scopo, prima è necessario introdurre il principio dell'utilità attesa.

1.3 Dalla teoria dell'utilità attesa al portafoglio ottimo

La teoria dell'utilità, sviluppata da Von Neumann e Morgenstern (1947), è un metodo che permette di selezionare la scelta di un consumatore razionale in condizioni di incertezza. Questo è un criterio molto simile a quello basato sulla massimizzazione del valor medio, però ci sono alcune differenze: innanzitutto bisogna evidenziare che il valore nominale di uno stesso determinato importo o rendimento ha un'importanza differente per ogni individuo; per di più è importante considerare da che punto di ricchezza si parte perché, nonostante maggiore ricchezza aumenti la "felicità" di qualsiasi individuo, l'utilità derivante dall'averne

un po' di più dipende dalla ricchezza posseduta inizialmente da quello specifico individuo. Quindi il criterio dell'utilità è di fondamentale importanza poiché tiene conto, da un punto di vista soggettivo, dell'importanza che ogni soggetto conferisce a un determinato investimento.

Dopo queste prime considerazioni, si rende necessario distinguere tra utilità degli importi certi e utilità attesa. La prima parte dall'idea che gli importi o i rendimenti futuri siano certi, quindi un individuo sceglierà l'investimento x_1 (che indica un importo certo) se la sua utilità è maggiore di quella dell'investimento x_2:

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

All'opposto, l'utilità attesa considera che l'individuo in oggetto si trovi all'istante *t*. Di conseguenza, gli oggetti che deve valutare al tempo *t* sono rischiosi, ossia si trova di fronte a più alternative senza sapere con certezza quale si verificherà (conoscendone, però, la distribuzione di probabilità). Quindi si calcola ancora un valor medio, però anziché valutare il valor medio delle realizzazioni, si calcola il valor medio delle utilità che vengono associate ad ogni realizzazione:

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

Come si può notare è una somma ponderata delle utilità di ciascun esito per la loro probabilità di realizzazione, considerando la variabile casuale come discreta.

La funzione dell'utilità attesa può essere vista come una funzione che associa ad ogni possibile scelta una corrispondente misura di utilità e può essere utilizzata per dare un ordinamento alle alternative rischiose: in pratica, ogni individuo sceglierà sempre la combinazione alla quale associa l'utilità attesa più elevata.

La costruzione di quest'ultima si fonda su alcuni assiomi che definiscono la logica sottostante al comportamento decisionale degli investitori. In particolare, tali proprietà, data una relazione binaria di preferenza debole totale $\succcurlyeq$, sono le seguenti.

- Riflessività: richiede un livello minimo di coerenza da parte dell'individuo, in quanto deve valutare ogni possibilità almeno tanto preferibile quanto sé stessa $x_i \succcurlyeq x_i$.

- Transitività: tale criterio di coerenza richiede che se un individuo deve scegliere tra tre alternative x_1 , x_2 e x_3 dove $x_1 \succcurlyeq x_2$ e $x_2 \succcurlyeq x_3$, allora deve essere vero che $x_1 \succcurlyeq x_3$.
- Completezza: si richiede che l'individuo sia sempre in grado di scegliere, cioè in presenza di due alternative sappia dare un giudizio di preferenza o di indifferenza tra le stesse.

Sono dati altri due assiomi ricavabili dal concetto di posizione finanziaria composta, ovvero quando, in presenza di un evento *A* con probabilità *a* (di valore strettamente compreso tra 0 e 1), l'individuo sceglierà l'alternativa x_1 se si verificherà l'evento *A*, mentre opterà per l'alternativa x_2 nell'eventualità opposta. Tale posizione composta rappresenta una variabile aleatoria mistura indicata come $x_1 a x_2$ e le proprietà sono le seguenti.

- Archimedea: quando in presenza di tre alternative $x_1\, x_2\, x_3$ dove $x_1 \preccurlyeq x_2 \preccurlyeq x_3$ allora $x_1 a x_3 \preccurlyeq x_2 \preccurlyeq x_1 b x_3$, con *a* e *b* comprese tra 0 e 1.
- Sostituzione: richiede che per qualunque probabilità *a,* se $x_1 \succcurlyeq x_2$, per ogni x_3 si deve avere $x_1 a x_3 \succcurlyeq x_2 a x_3$.

Queste ultime due proprietà implicano altre tre proprietà più deboli ma comunque molto importanti e sono: la proprietà di continuità, la proprietà di monotonia e la proprietà di consistenza.
Pertanto, se una relazione binaria di preferenza gode dei primi cinque assiomi e gli individui agiscono in modo razionale, è possibile creare una funzione di utilità che rappresenta la relazione stessa, secondo la quale l'alternativa x_1 è preferita all'alternativa x_2 (entrambe variabili casuali discrete) se e solo se l'utilità attesa di x_1 è superiore a quella di x_2 ; inoltre la funzione di utilità è unica a meno di una trasformazione lineare crescente.
Gli assiomi definiti precedentemente sono anche la causa delle critiche riferite alla teoria dell'utilità. Difatti, molti economisti non sono d'accordo con il presupposto della teoria che definisce gli individui come razionali e propensi ad adottare un comportamento che rispetta sempre i cinque assiomi.
Per quanto riguarda le caratteristiche che possiede una funzione di utilità, queste sono in particolare due:

- Considerando che gli individui sono "ingordi", ossia massimizzano il profitto, più elevato è l'importo (anche se più rischioso) maggiore è l'utilità. Quindi, se la funzione di utilità è derivabile, la sua derivata prima, detta utilità marginale, sarà strettamente positiva su tutto il dominio.
- La seconda caratteristica recepisce il fatto che gli individui sono avversi al rischio. A seconda della ricchezza di partenza dell'individuo, l'aumento di utilità c'è sempre, anche se in misura minore. Questo implica la concavità della funzione di utilità, in quanto aumenta meno proporzionalmente all'aumentare del capitale, dunque la sua derivata seconda sarà strettamente negativa.

Affermare che tutti gli individui sono avversi al rischio è coerente con la teoria economica, che definisce un investitore come un soggetto razionale. Questo non è necessariamente vero nella realtà, dove alcuni individui possono avere:

- Una funzione di utilità lineare se sono indifferenti al rischio, ovvero sono indifferenti tra un'alternativa di incertezza e quella di certezza.
- Una funzione di utilità convessa se invece sono propensi al rischio, preferendo un'alternativa con risultato rischioso rispetto alla stessa con risultato certo.

Comunque definire un soggetto come "avverso al rischio" non significa che questo non intenda assumere rischi di nessun tipo, significa che è disposto ad accettare un determinato rendimento con il minor rischio possibile.
Misurare l'avversione al rischio di un determinato individuo è molto complesso, perché dipende da una molteplicità di fattori tra cui: il livello di percezione del rischio, la situazione finanziaria, l'età e la prospettiva di reddito futuro. In genere, per fare ciò viene utilizzato un questionario che restituisce una valutazione qualitativa dell'individuo.
In letteratura si possono trovare molte funzioni di utilità, ad esempio quella esponenziale o quella logaritmica, tuttavia nella moderna teoria del portafoglio la scelta ricade sulla funzione di utilità quadratica:

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

Essendo una parabola può essere utilizzata solo per la sua parte crescente, ossia il ramo ascendente che va da meno infinito al punto *1/a*, proprio perché l'utilità non decresce mai (vedi Figura 4).

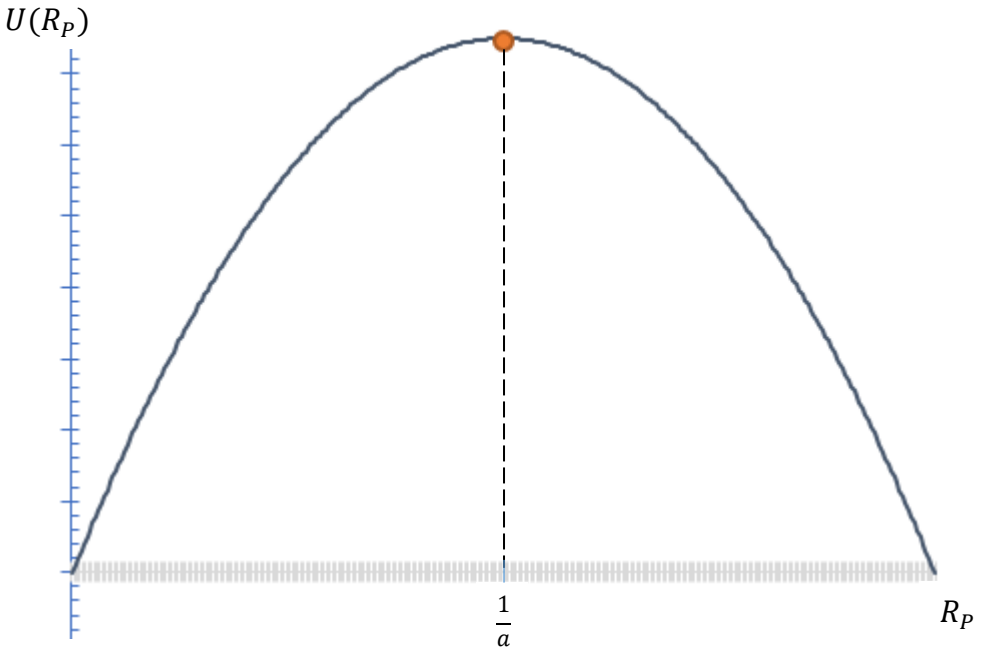


Figura 4: Funzione di utilità attesa.

A differenza della funzione logaritmica, questa ha il vantaggio di poter essere utilizzata anche per importi negativi, cioè per misurare la disutilità delle perdite.
All'interno della formula è presente il parametro *a* che indica l'avversione al rischio di un determinato individuo.
Calcolando il valore atteso di questa funzione di utilità, si può apprezzare come dipenda solamente dalla media e dalla varianza:

$$E[U(R_P)] = E(R_P) - \frac{1}{2}aE(R_P^2) = \mu - \frac{1}{2}a(\sigma^2 + \mu^2)$$

in quanto $\sigma^2 = E\left(R_p^2\right) - [E(R_P)]^2$.

È possibile dimostrare che la coerenza tra il criterio media-varianza e la funzione di utilità attesa è garantita solo in presenza di una delle seguenti condizioni:

- quando la funzione di utilità attesa è quadratica;

- quando la distribuzione di probabilità congiunta dei rendimenti è ellittica[6].

La curva d'indifferenza[7] della funzione di utilità quadratica può essere ricavata dall'equazione dell'utilità attesa:

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

dove U^* è un prefissato valore di utilità attesa.

Le curve d'indifferenza sono crescenti (rispetto a μ) poiché la loro derivata prima è positiva rispetto a σ:

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

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

quindi il "prezzo da pagare" per mantenere l'utilità costante all'aumentare del rischio è un maggior rendimento. Inoltre la positività della derivata seconda garantisce la convessità delle stesse:

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

Ovviamente deve valere il vincolo di non negatività $\mu < \frac{1}{a}$. Nella Figura 5 sono raffigurate alcune curve d'indifferenza.

[6] Si definisce una distribuzione come ellittica quando tagliandola in qualsiasi punto si ottiene un'elisse.

[7] Le curve d'indifferenza possono essere definite come tutte le combinazioni rischio-rendimento che presentano la stessa utilità attesa per un determinato investitore. In altri termini, date due scelte d'investimento appartenenti alla stessa curva d'indifferenza, per un investitore sarà indifferente mettere tutto il suo capitale nella prima o nella seconda scelta.

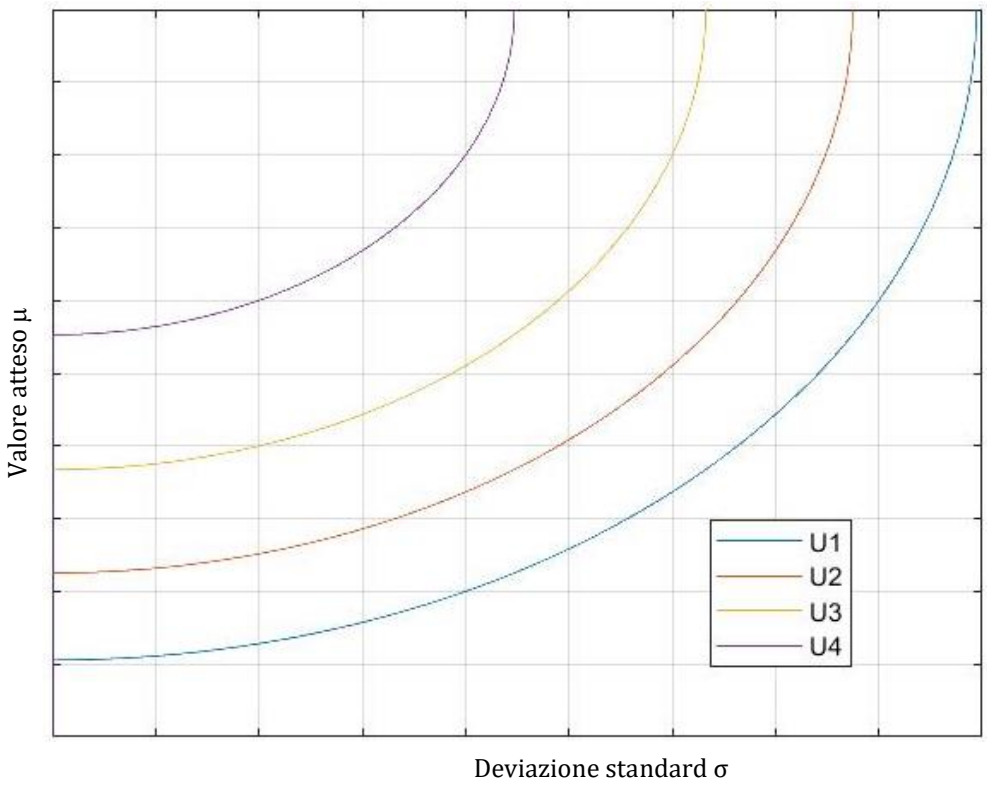


Figura 5: Curve d'indifferenza.

Collocando la frontiera efficiente sullo stesso piano delle curve d'indifferenza si può concludere il processo di selezione alla Markowitz. Infatti, il portafoglio ottimo sarà quello corrispondente al punto di tangenza tra le due curve, ossia quel portafoglio che soddisfa la duplice condizione di posizionarsi sulla frontiera efficiente e di massimizzare l'utilità attesa.

1.4 Critiche al modello di selezione alla Markowitz

Negli anni '50, la brillante intuizione di Markowitz fu immediatamente considerata rivoluzionaria nel mondo finanziario poiché, prima della pubblicazione del suo celebre *paper*, i portafogli azionari erano costruiti in maniera *naive*. Il suo apporto può essere considerato la nascita della finanza moderna.

Tuttavia, tale modello presenta una serie di limiti che emergono in particolar modo dalla sua applicazione pratica: da un lato riguardano le ipotesi di base e dall'altro lato i vincoli utilizzati.

Uno dei primi limiti venne messo in evidenza da Sharpe, partendo dal presupposto che il numero di input richiesto dal modello di selezione di Markowitz è molto

elevato, in quanto necessita del calcolo di *N* medie dei rendimenti, *N* varianze dei rendimenti e *N(N-1)/2* covarianze tra i rendimenti. Per cui, a titolo esemplificativo, un semplice portafoglio composto da 200 titoli richiede la stima di oltre 40000 parametri.
Sharpe osservò che la stima di un numero così elevato di parametri è problematica per almeno due motivi:

- il primo, molto più evidente negli anni cinquanta del secolo scorso, è il tempo di calcolo necessario per effettuare tutte le stime;
- il secondo deriva dal fatto che non si conoscono i veri parametri, di conseguenza bisogna effettuare delle stime puntuali che saranno comunque affette da un errore di stima.

Con l'obbiettivo di mitigare tali problemi, Sharpe sviluppò un modello, detto diagonale, che permette di ridurre il numero di parametri da stimare, consentendo di beneficiare di un risparmio in termini di tempo di calcolo e di numero di errori di stima.
Una seconda critica può essere attribuita all'instabilità nel tempo delle soluzioni *mean-variance*: i processi che generano i rendimenti sono non stazionari nel tempo, cioè le medie, le varianze e le covarianze future non sono uguali a quelle passate. Di conseguenza, si possono ottenere frontiere efficienti differenti, anche solo considerando intervalli temporali consecutivi. I modelli di ribilanciamento hanno l'obbiettivo di porre rimedio a tale problema.
Continuando la rassegna dei principali limiti del modello di Markowitz, a questo punto l'interesse ricade sulle due condizioni che garantiscono la coerenza tra il criterio media-varianza e la teoria dell'utilità attesa definiti precedentemente.
La prima assunzione, ossia la ellitticità della funzione congiunta di densità di probabilità dei rendimenti degli *asset*, fonda le sue radici nel lontano 1900, quando il matematico L. Bachelier, partendo da un modello che rappresenta gli investitori come delle particelle che si incontrano casualmente e si scambiano informazioni, arrivò alla conclusione deduttiva che la distribuzione dei rendimenti logaritmici debba essere normale, ovvero un caso particolare di distribuzione ellittica. Tuttavia, è stato mostrato empiricamente che questo non è sempre vero nella realtà. Uno tra i primi ad accorgersene fu Mandelbrot nel 1963, mostrando che le distribuzioni dei

rendimenti logaritmici dei prezzi azionari presentano delle code più spesse rispetto alla normale.
A riprova di ciò, uno dei fatti stilizzati osservati da Cont (2001) è che le distribuzioni dei rendimenti mostrano:

- asimmetria negativa, ossia presentano una coda allungata a sinistra, questo vuol dire che le perdite estreme possono essere molto più elevate dei profitti che si trovano nell'estremità opposta della distribuzione;
- leptocurtosi, cioè presentano una curtosi elevata che implica una distribuzione più appuntita rispetto alla normale.

È evidente che nel momento in cui si analizzano distribuzioni non normali sarebbe estremamente riduttivo limitarsi al calcolo dei primi due momenti.
Per quanto riguarda il calcolo dell'asimmetria, nella letteratura si possono trovare molte misure presentate nel corso degli anni. Il primo a proporre una misura di questo tipo fu Pearson nel 1895:

$$S = \frac{\mu - m}{\sigma},$$

dove μ e m rappresentano rispettivamente la media e la moda della distribuzione, mentre σ è la deviazione standard.
Un metodo alternativo utilizzato da Peirò (1999) è quello di specificare due sottoinsiemi del campione oggetto di analisi:

$$R^+ = \{r_t - \mu | r_t > \mu\}$$
$$|R^-| = \{\mu - r_t | r_t < \mu\}$$

i quali definiscono rispettivamente i rendimenti in eccesso positivi e il valore assoluto dei rendimenti in eccesso negativi. Se tali sottoinsiemi non presentano la stessa distribuzione allora i rendimenti non sono simmetrici rispetto alla media.
L'indice di asimmetria campionario più utilizzato in letteratura è dato da

$$\hat{S} = \frac{\sum_{t=1}^{T} (r_t - \mu)^3 / T}{\hat{\sigma}^3}$$

dove

- T è il numero di osservazioni,
- r_t è il rendimento al tempo t,
- μ è la media campionaria,
- $\hat{\sigma}$ è la deviazione standard campionaria.

Dato che la distribuzione normale è simmetrica rispetto alla media, essa presenterà un valore di tale indice sempre pari a zero.

La distribuzione asintotica di $\hat{S}$, sotto l'ipotesi di normalità, è data da $\hat{S} \to N(0, \frac{6}{T})$.

Concentrandosi ora sulla curtosi, detta anche momento quarto, questa può essere interpretata come una misura che quantifica la distanza di una distribuzione empirica da una distribuzione gaussiana a causa delle cosiddette *fat thails*. L'indice campionario di curtosi è dato da:

$$\hat{k} = \frac{\sum_{t=1}^{T} \frac{(r_t - \mu)^4}{T}}{\hat{\sigma}^4}$$

La distribuzione normale presenta un valore di tale indice pari a 3.

Riprendendo quanto detto ad inizio paragrafo è stato mostrato empiricamente che le distribuzioni dei rendimenti dei titoli azionari si allontanano da una normale a causa sia della non nullità del momento terzo, sia di un valore del momento quarto maggiore di 3. Da questo ne deriva che gli eventi estremi sono più probabili di quanto non indichi la normale, poiché quest'ultima presenta una minor area al di sotto delle code. Pertanto, se un investitore decide di considerare normali i rendimenti dei titoli per investire in un portafoglio, rischia di sottostimare pesantemente gli eventi estremi (in particolare quelli negativi), ossia il rischio di incorrere in perdite elevate. A titolo esemplificativo, basti pensare al crollo dei contratti *futures* con sottostante l'indice *SP500* avvenuto nell'ottobre del 1987: ipotizzando che le fluttuazioni di borsa seguissero una distribuzione normale, a tale ribasso fu attribuita una probabilità di 1 su 10^{160} di verificarsi. Discorso analogo vale per la drammatica crisi del 2008, dove si verificarono una serie di eventi che "sarebbero dovuti accadere solo una volta ogni 100000 anni circa"[8].

Inoltre, come evidenziato da Szegö (2002) e Cont (2001), la covarianza non è un

[8] Bonner B. su MoneyWeek, 25 *Standard Deviations in a Blue Moon*, 13 novembre 2007.

buon indicatore di dipendenza in presenza di distribuzioni non ellittiche, in quanto non riesce a misurare la correlazione degli eventi estremi negativi: la dipendenza è molto più elevata sulla coda sinistra rispetto a quella di destra. In genere, per superare tale ostacolo, vengono utilizzate le funzioni copula[9].

Con il fine di modellizzare in modo più adeguato le distribuzioni empiriche delle variazioni dei prezzi dei titoli azionari, nel corso degli anni sono state proposte un numero notevole di distribuzioni alternative, al punto che Cont (2001) affermò *«Fitting various functional forms to the distribution of stock returns and stock price changes has become a popular pastime».*

In tale ambito, Mandelbrot (1963) e Fama (1963) suggerirono di impiegare la famiglia delle distribuzioni Lévy-Pareto stabili sviluppate da P. Lévy nel 1925 e considerate tutt'ora una delle principali concorrenti della distribuzione normale. Queste presentano una funzione caratteristica data da:

$$E\left(e^{i\vartheta X}\right) = \begin{cases} exp\left\{-\sigma^{\alpha}|\vartheta|^{\alpha}\left(1 - i\beta sgn(\vartheta)tg\left(\frac{\alpha\pi}{2}\right)\right) + i\mu\vartheta\right\} \; se\; \alpha \neq 1 \\ exp\left\{-\sigma|\vartheta|\left(1 - i\beta\frac{2}{\pi}sgn(\vartheta)\log(|\vartheta|)\right) + i\mu\vartheta\right\} \; se\; \alpha = 1 \end{cases}$$

dove $i = \sqrt{-1}$.

Stimare i quattro parametri $\alpha \in (0,2]$, $\beta \in [-1,1]$, $\mu \in R$ e $\sigma \in [0, +\infty)$ non è una cosa agevole. In particolare, il parametro α (detto anche caratteristico o indice di stabilità) è un valore legato alla curtosi della distribuzione di probabilità e fornisce una misura dell'area sottesa dalle code. Quando è pari a 2 si ottiene esattamente la distribuzione normale con media nulla e varianza $\sigma^2 = 2c$. Man mano che α decresce da due verso zero accade che la distribuzione diventa sempre più leptocurtica e le code sempre più grosse. Da questo ne deriva che modellizzare tutti i rendimenti con la normale vuol dire scegliere, tra tutte le possibili Lévy-Pareto stabili, quella per il quale gli eventi estremi sono i meno rischiosi in assoluto. Un'altra particolarità fondamentale è che quando α è minore di 2, la varianza tende a più infinito poiché l'integrale improprio utilizzato per calcolarla è finito solo se α è uguale a 2. Questo non vuol dire che il rischio diventa infinito, ma solo che la

[9] Per maggiori informazioni si veda *An introduction to Copulas,* Nelsen R.B., Springer Science & Business Media, 1999.

varianza non è più un indicatore corretto per misurarlo. Da un punto di vista empirico, in genere i mercati finanziari presentano un α che appartiene ad un intervallo compreso tra 1.5 e 2.
Il parametro β misura il livello di asimmetria della distribuzione:

- se $\beta < 0$ indica asimmetria a sinistra rispetto a μ;
- se $\beta = 0$ indica simmetria rispetto a μ;
- se $\beta > 0$ indica asimmetria a destra rispetto a μ.

Per quanto riguarda il parametro di locazione μ, questo è un indice di posizione della distribuzione di probabilità: coincide con il valor medio se α è compreso fra 1 e 2. Dato che anche questo si calcola mediante l'integrale generalizzato, si può dimostrare che quando α è minore di uno la media tende a infinito e quindi non è più un buon metodo per misurare la performance.
Infine σ è un parametro di variabilità e può essere interpretato come la misura di rischio per le Lèvy-Pareto stabili.
Spostando ora l'attenzione sulla seconda condizione che garantisce la coerenza tra il criterio media-varianza e la teoria dell'utilità attesa, ovvero l'assunzione che l'investitore presenti una funzione di utilità quadratica, pure in questo caso ci si scontra con un insieme di problemi non indifferenti. Come evidenziato nel paragrafo precedente, tale funzione è una parabola concava verso il basso quindi, affinché sia rispettato l'assioma di non sazietà, può essere utilizzata solo fino al punto in cui la sua derivata prima è pari a zero. Pertanto se i rendimenti possibili superassero il punto di massimo della funzione, questa inizierebbe a decrescere interpretando un aumento della ricchezza come un evento negativo che riduce l'utilità dell'individuo. Inoltre appare evidente come sia riduttivo modellizzare l'utilità di un individuo mediante una funzione quadratica che, da un lato non dipende da alcun momento superiore al secondo, dall'altro lato riassume tutta l'informazione relativa all'avversione al rischio di un investitore semplicemente con il parametro α. Più in dettaglio, l'utilizzo di tale funzione in presenza di distribuzioni non normali porta a considerare un qualsiasi individuo come indifferente a qualsiasi momento superiore alla varianza. Al contrario, studi come quello di Conine e Tamarkin (1981), hanno dimostrato come gli investitori siano propensi a "sacrificare" parte dei potenziali vantaggi derivanti dalla diversificazione, a favore di una maggiore asimmetria

positiva. Quindi sono disposti a detenere portafogli non efficienti sul piano *mean-variance*, ma efficienti su quello *mean-variance-skewness*. Alderfer e Bierman (1970), attraverso un'indagine empirica dove posero gli investitori di fronte a differenti strategie di investimento caratterizzate da performance e rischio simili ma indici di asimmetria diversi, provarono che la preferenza dei partecipanti ricadeva su quelle strategie che presentavano una distribuzione asimmetrica a destra.

La ragionevolezza di questa preferenza è giustificata dal fatto che un'asimmetria positiva implica, in termini finanziari, una maggiore probabilità di ottenere performance maggiori della media.

Più in generale, la preferenza di un investitore è positiva (negativa) per valori positivi (negativi) di ogni momento centrale dispari e negativa per ogni momento centrale pari (Scott e Horvath, 1980).

Un'altra evidenza empirica, che sfavorisce la diversificazione, deriva dal fatto che in periodi di forte crisi dei mercati finanziari si verifica un aumento della covarianza anche tra quei titoli che non appartengono allo stesso settore: ad esempio, tale evento si verificò durante le crisi del 2008 scoppiata in seguito al fallimento della banca d'investimento Lehman Brothers. Indici di covarianza elevati di fatto eliminano i vantaggi che favoreggiano una maggiore diversificazione. Tuttavia, data la formula della covarianza $C_{ij} = \sigma_i \sigma_j \rho_{ij}$, si rende necessario puntualizzare che una variazione di quest'ultima potrebbe non essere dovuta ad un aumento o diminuzione di ρ_{ij}, ma semplicemente dalla variazione della volatilità dovuta alla natura eteroschedastica delle serie storiche dei rendimenti dei titoli azionari.

I limiti definiti fino a questo punto riguardano in particolare la lontananza dalla realtà delle ipotesi sottostanti al criterio di selezione proposto da Markowitz. Ora verrà presentato un secondo insieme di criticità legate ai vincoli e alla funzione obbiettivo del processo di ottimizzazione.

Il modello di programmazione quadratica lineare iniziale, sviluppato per ricavare la frontiera efficiente, presuppone che le vendite allo scoperto siano sempre permesse, poiché considera solamente due vincoli: il vincolo di bilancio e il vincolo relativo al rendimento desiderato. Tale assunzione, ossia che sia sempre possibile prendere a prestito titoli con il fine di venderli ad altri soggetti per *speculare* sul loro ribasso, non è sempre vera nella realtà. Non mancano esempi storici dove le autorità di vigilanza decisero di vietare questa operazione: in Italia il caso più recente risale

solo al novembre del 2017, quando in seguito ad una perdita superiore al 20%, la CONSOB decise di vietare temporaneamente le vendite allo scoperto sui titoli del Credito Valtellinese. Per di più, gli elevati costi e commissioni associati alla vendita allo scoperto di titoli intaccano in maniera rilevante i potenziali profitti, scoraggiando ulteriormente l'utilizzo di tale alternativa da parte degli investitori.
Da queste osservazioni ne deriva la prima importante estensione del modello classico di Markowitz:

$$min_{x_1,\dots,x_n}\ x'Vx$$
$$s.t.\ \begin{cases} x'r = \pi \\ x'e = 1 \\ x_i \geq 0\ \forall i \end{cases}$$

dove $x_i \geq 0\ \forall i$ è il vincolo di non negatività sui pesi, ossia il divieto di andare *short*.
Sebbene questa estensione attribuisca maggiore validità pratica al problema di partenza, rimangono ulteriori aspetti fondamentali che non vengono considerati, in particolare:

- La presenza di aliquote di tassazione, *fees* e costi di transazione. Dato che i costi aumentano all'aumentare delle transazioni effettuate, l'aggiunta di tali vincoli al modello frenano la diversificazione.
- La presenza del vincolo legato al fatto che un determinato titolo può essere acquistato o venduto solamente in un numero intero di lotti minimi di transazione.

L'ultima critica ricade sulla funzione obiettivo impiegata nel modello, ossia la misura di rischio. Minimizzare la varianza significa rendere minima la media delle distanze delle varie realizzazioni dal loro valore atteso. Quindi non cattura l'avversione al rischio di un determinato investitore e considera indistintamente rischiose sia le *sovra-performance* che le *sotto-performance*. È intuitivo pensare che nessun investitore razionale considererebbe un rischio la possibilità di ottenere profitti sopra la media, tanto meno vorrebbe minimizzare tale eventualità. Inoltre, l'utilizzo di questa misura di dispersione, per distribuzioni che presentano asimmetria negativa o positiva, porta rispettivamente a sottostimare il rischio o minimizzare la possibilità di ottenere guadagni elevati.

Lo stesso Markowitz, con l'obbiettivo di superare i problemi legati alla varianza, propose una misura di rischio alternativa, detta semi-varianza, che considera solamente il cosiddetto *downside risk*, ovvero il rischio legato ai rendimenti al di sotto di un prefissato *benchmark* (in genere zero o la media):

$$Semi - Var(r) = \frac{\sum_{t=1, r_t<\bar{r}}^{T}(r_t - \bar{r})^2}{T}.$$

In altre parole, questa misura di rischio considera solamente il lato sinistro della distribuzione a partire da un determinato valore. Il problema principale di questa misura è la sua scarsa trattabilità da una prospettiva analitica.
Una terza misura di rischio proposta, e sempre legata al concetto di dispersione, è la *mean absolute-deviation*:

$$MAD(r) = \frac{\sum_{t=1}^{T}|r_t - \bar{r}|}{T}$$

La presenza del valore assoluto rende anche questa misura complicata da trattare analiticamente. Inoltre, nonostante sia meno sensibile agli *outliers*[10], presenta gli stessi problemi della varianza.
Nel prossimo capitolo verranno presentate dettagliatamente le principali misure di rischio proposte recentemente.
Si rende necessario evidenziare fin da subito che le attività finanziarie considerate in questo lavoro non saranno titoli azionari, ma nove contratti future che differiscono tra loro per la tipologia di sottostante.

[10] Osservazioni anomale rispetto alle altre disponibili.

CAPITOLO 2

DALLE MISURE DI RISCHIO PREMODERNE ALLE MISURE DI RISCHIO COERENTI

In questo secondo capitolo, prima verrà illustrata la principale misura di rischio premoderna utilizzata nella pratica finanziaria, ovvero il Value at Risk, per poi spostare l'attenzione verso le quattro proprietà che una generica misura di rischio deve possedere per essere definita coerente. Infine, dopo aver presentato le principali misure di rischio coerenti proposte in letteratura, si introdurrà la nuova classe delle misure di rischio Two-sided.

2.1 Il Value at Risk

Il Value at Risk[11] può essere definito come una misura di rischio *premoderna*, in quanto si trova in una posizione intermedia tra le misure di rischio classiche, definite nel precedente capitolo, e le misure di rischio moderne, che verranno presentante nel seguito.

Prima di fornire una definizione puntuale del VaR è indispensabile ripercorrere le fasi storiche che hanno portato al suo sviluppo e utilizzo.

Alla fine degli anni '80, non esisteva una misura di rischio unica che riuscisse a catturare il rischio complessivo di un portafoglio o di un'attività finanziaria complessa. In particolare, la moltitudine di indicatori di rischio sviluppati fino a quel momento erano tutti legati ad uno specifico strumento, ad esempio:

- per quanto riguarda le obbligazioni, il rischio di tasso veniva misurato principalmente mediante la durata media finanziaria o *duration*, ossia un indicatore che fornisce la sensibilità del prezzo di un titolo in seguito alla variazione del tasso d'interesse;

[11] VaR d'ora in poi.

- per misurare il rischio associato ai titoli azionari si sfruttava il coefficiente β, definito all'interno del celebre modello del *Capital Asset Pricing Model*;
- il rischio associato alle opzioni veniva valutato attraverso le *greche*, le quali determinano quanto varia il prezzo di un'opzione al variare di un insieme di fattori che influenzano la stessa;
- per quanto riguarda il rischio di credito, la misura maggiormente utilizzata era il *rating*, ovvero un voto sulla solvibilità e solidità di una determinata società, espresso in lettere e fornito da agenzie legalmente autorizzate a tale scopo.

L'incomparabilità di tali misure rendeva impossibile definire un indicatore completo che considerasse in modo esaustivo tutti i rischi specifici. In particolare, non esisteva una misura di rischio che permettesse di riassumere in un singolo valore ogni rischio che compone la classe dei rischi di mercato[12].

Questi problemi erano particolarmente evidenti a Dennis Weatherstone, a quel tempo amministratore delegato della banca d'investimento americana J.P. Morgan. Durante una riunione chiese ai suoi dirigenti quale fosse il rischio complessivo a cui è soggetta la banca, ricevendo, tuttavia, risposte differenti relativamente all'area a cui appartenevano gli stessi. Di conseguenza, trasferì tale richiesta ai suoi migliori analisti, i quali si misero subito all'opera ed implementarono quello che può essere considerato il *Sacro Graal* tra le misure di rischio, ossia il VaR.

Tutti i giorni, entro le 16:15, un documento contenente il valore di tale misura di rischio "complessiva" doveva essere preparato e consegnato a Weatherstone.

Nel 1994, i dirigenti di J.P. Morgan decisero di pubblicare il documento *RiskMetrics*[13], mediante il quale resero noto, oltre le pareti della banca, il VaR. Da quel giorno in poi, ci fu un rapido sviluppo ed una espansione dei metodi e dei modelli per la quantificazione dei rischi.

Il notevole successo, che contraddistinse fin da subito il VaR, è da impuntare principalmente alla sua definizione di immediata comprensione: dato un certo

[12] Rischio relativo agli effetti imprevisti sul valore di mercato di attività e passività prodotti da variazioni dei tassi di interesse, dei tassi di cambio e da altri prezzi delle attività. È una tipologia di rischio che sorge dalla detenzione di differenti strumenti finanziari.

[13] Per maggiori informazioni si veda J.P. MORGAN, $RiskMetrics^{TM}$ – *Technical Document, Fourth Edition (1996).*

livello di confidenza $\alpha \in [0; 1]$, il VaR è la massima perdita potenziale a cui un determinato portafoglio potrebbe andare incontro nell' $\alpha\%$ dei casi e in un definito intervallo di tempo (detto *holding period*) fissato a priori. In altri termini, non è altro che il valore massimo che un individuo può attendersi di perdere con una determinata probabilità e in orizzonte temporale che può essere un giorno, una settimana, un anno, ... Risulta evidente che un aumento di quest'ultimo comporterà una crescita del valore del VaR.

È prassi utilizzare un intervallo di confidenza α superiore o uguale al 95% e un arco temporale che dipende dalla liquidità dell'attività oggetto di analisi.

Nonostante la chiarezza delle definizioni presentate nelle righe precedenti, solamente attraverso un esempio si può apprezzare a pieno la semplicità concettuale di questa misura di rischio: se un portafoglio composto da titoli azionari presenta un VaR di 100€ con $\alpha = 95\%$ e $t = 5$ giorni, significa che con una probabilità del 95% il portafoglio in questione potrebbe subire una perdita massima di 100€ nei successivi 5 giorni; perciò, la probabilità di superare tale soglia è pari al 5%.

La Figura 6 esemplifica ulteriormente quanto detto, evidenziando come la misura di rischio in oggetto sia pari a quel valore che presenta alla sua sinistra il 5% di probabilità ($1- \alpha\%$) e alla sua destra il rimanente 95% ($\alpha\%$).

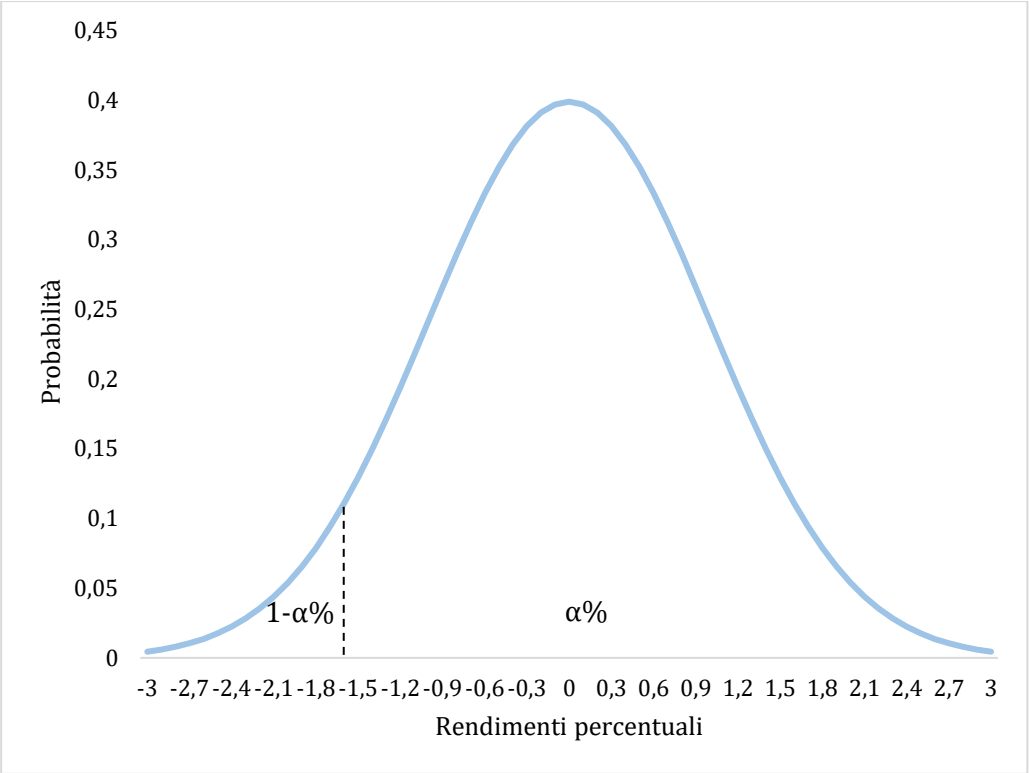


Figura 6: Value at risk con intervallo di confidenza pari al 95%.

Più in dettaglio, da un punto di vista probabilistico, il VaR soddisfa la seguente uguaglianza:

$$P(Loss > VaR_\alpha) = 1 - \alpha,$$

dove *Loss* rappresenta una generica variabile casuale che rappresenta le perdite.

Come si può notare, il VaR è principalmente espresso in termini monetari, e non percentuali, pertanto, passando momentaneamente ad un'ottica di gestione del rischio, può essere interpretato come il capitale da accantonare per far fronte alle future perdite massime potenziali con una probabilità pari ad α. Di conseguenza si genera un trade off non indifferente poiché, abbassando il valore di α, da un lato sarà possibile accantonare meno capitale, ma dall'altro lato si considererà un evento che può capitare un numero maggiore di volte.
Nei seguenti paragrafi prima verranno presentati in modo sintetico i tre metodi classici che possono essere utilizzati per stimare il VaR, poi le critiche connesse a questa misura di rischio.

2.1.1 Metodi di stima del Value at Risk

La computazione del VaR può essere eseguita seguendo 3 processi differenti:

- il metodo parametrico;
- il metodo storico;
- la simulazione Monte Carlo.

Il primo approccio, detto anche metodo "varianza-covarianza", è senza dubbio il più agevole da implementare, nonché il più utilizzato (rispetto agli altri due metodi).
Tale metodo si fonda sull'ipotesi di *normalità* della distribuzione di probabilità dei profitti e delle perdite del portafoglio oggetto di analisi. Più dettagliatamente, si assume che i rendimenti dei titoli che compongono il portafoglio siano funzioni lineari di un insieme di fattori di rischio che seguono a loro volta una distribuzione Gaussiana: pertanto, la media e la varianza dei rendimenti delle distribuzioni dei vari titoli possono essere calcolate partendo dalla media e dalla varianza dei fattori di rischio relativi ad ogni attività.
Dato che la distribuzione normale è simmetrica, dalla statistica si può ricavare che un determinato valore percentuale delle osservazioni si trova sempre all'interno di due soglie che corrispondono alla differenza/somma tra la media μ e il prodotto di un dato valore α per la deviazione standard σ:

- $\mu \pm 1\sigma = 68\%$
- $\mu \pm 1.96\sigma = 95\%$
- $\mu \pm 2.58\sigma = 99\%$.

A questo punto può essere definita la formula del VaR considerando inizialmente una sola attività:

$$VaR = V \times \alpha \times \sigma_{gg} \times \sqrt{gg},$$

dove

- V è il valore dell'esposizione;
- α è il valore che indica per quante volte deve essere moltiplicata la deviazione standard per conseguire il livello di confidenza desiderato;
- σ_{gg} è la deviazione standard del rendimento del titolo;
- gg è il numero di giorni adottati per la valutazione.

Il calcolo del VaR di un portafoglio composto da n attività finanziarie risulta più complesso, poiché si rende necessario determinare anche le correlazioni tra i singoli titoli che lo compongono. Utilizzando la notazione matriciale, in questo caso può essere formulato come:

$$VaR_P = \sqrt{VaR \times CORR \times VaR^T},$$

dove VaR è il vettore contenente i singoli VaR dei titoli che compongono il portafoglio e $CORR$ è la matrice delle correlazioni tra le diverse attività presenti nel portafoglio.

Il limite principale del metodo parametrico riguarda l'ipotesi di normalità delle distribuzioni dei rendimenti che, come visto nel capitolo precedente, è lontana dall'evidenza empirica. Inoltre, il suo utilizzo ideale ricade principalmente su quei portafogli composti da titoli che presentano *payoff* lineari come le azioni.

Il secondo metodo, ovvero la simulazione storica, si differenzia dal precedente in quanto non richiede né l'assunzione a priori di una distribuzione di probabilità, né la stima della matrice delle correlazioni.

Partendo dalla serie storica delle perdite e dei profitti del portafoglio, il procedimento di tale approccio può essere sintetizzato in tre fasi:

- disporre tutti i valori della serie storica in maniera decrescente, ossia dal maggior profitto alla perdita più negativa;
- applicazione della formula $[(n° \, osservazioni + 1) \times intervallo \, di \, confidenza]$, mediante la quale si ottiene la posizione, all'interno della serie formata precedentemente, dove si trova il valore corrispondente al VaR.

Come si può vedere, un'ipotesi implicita connessa a questo metodo è la *stazionarietà* dei processi che generano le serie storiche: difatti l'utilizzo delle osservazioni passate, con il fine di stimare il valore a rischio, comporta l'assunzione che le distribuzioni di quest'ultime siano perfettamente identiche a quelle future. A livello empirico la presenza di eteroschedasticità nelle serie storiche dei titoli azionari rende tale ipotesi il principale limite del metodo storico.

Il terzo ed ultimo approccio, ossia la simulazione Monte Carlo, è il può complesso da

definire ed implementare. In modo analogo alla simulazione storica, si fonda sulle osservazioni passate dei rendimenti del portafoglio oggetto di analisi ed è utilizzato essenzialmente in presenza di titoli con *payoff* non lineare (come le opzioni).

La prima fase di questo processo consiste nel duplice compito di individuare i fattori di rischio, e determinare le formule matematiche che definiscono come questi fattori influenzano il valore dei singoli strumenti finanziari che compongono il portafoglio. Passando alla seconda fase, questa necessita l'individuazione a priori della distribuzione di probabilità, e dei relativi parametri per i fattori di rischio selezionati: mentre nella simulazione storica si assume che le distribuzioni future siano uguali a quelle passate, in questo caso sarà l'analista ad identificare il processo stocastico che meglio si adatterà alle osservazioni future (che comunque terrà conto di quelle passate). Successivamente si generano un numero *N* di valori casuali, con lo scopo di ottenere *N* valori ipotetici dei fattori di rischio da cui si ricaveranno *N* valori ipotetici del portafoglio.[14] Infine, l'ultimo passo della procedura è lo stesso del metodo precedente, ossia si ordinano in modo crescente le perdite e i profitti ipotetici ottenuti, andando poi a selezionare il valore corrispondente al VaR.

Il modello stocastico adottato per modellare i fattori di rischio è indubbiamente il punto critico della simulazione Monte Carlo: una scelta poco realistica comporterebbe una perdita di significato del VaR come misura di rischio. Uno dei modelli più utilizzati a tale scopo è il moto browniano geometrico[15].

2.1.2 Limiti del Value at Risk

Gli aspetti positivi del VaR, negli anni immediatamente successivi al suo sviluppo, erano particolarmente evidenti alle autorità di vigilanza, al punto che decisero di imporre il suo utilizzo agli intermediari finanziari con il fine di determinare il capitale da accantonare a riserva per fronteggiare il rischio di mercato a cui erano sottoposti.

Tuttavia, nella letteratura successiva, diversi autori non tardarono ad accorgersi dei problemi non indifferenti legati all'utilizzo di tale indicatore come misura di rischio.

[14] All'aumentare di *N* la stima converge al valore reale, tuttavia la procedura richiederà un maggiore tempo-macchina.

[15] È un processo stocastico in tempo continuo dove il logaritmo del valore variabile nel tempo segue un processo di Wiener. In particolare, è utilizzato nell'ambito del pricing delle opzioni.

Un primo evidente difetto del VaR può essere agevolmente desunto dalla sua definizione originale: dato che individua la massima perdita potenziale all'interno di un intervallo di confidenza $\alpha\%$, non fornisce alcuna informazione riguardo alle perdite che si trovano al di là di tale soglia. In altri termini, essendo una misura di tipo percentile, non considera il lato della coda sinistra della distribuzione che presenta le perdite più estreme rispetto al livello del VaR stesso. Questo limite conduce a risultati troppo ottimisti, soprattutto in presenza di distribuzioni che presentano una forte asimmetria negativa.

A titolo esemplificativo, si osservi la Figura 7 e la Figura 8: entrambe rappresentano le distribuzioni di probabilità dei rendimenti di due differenti portafogli A e B.

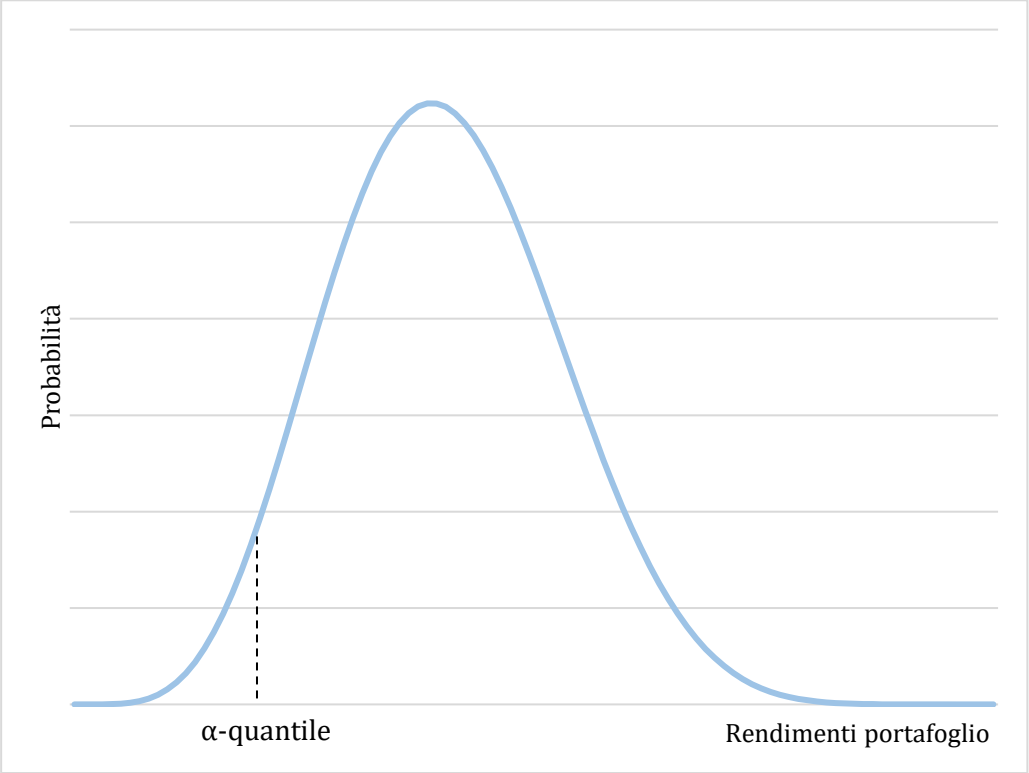


Figura 7: distribuzione dei rendimenti del portafoglio A.

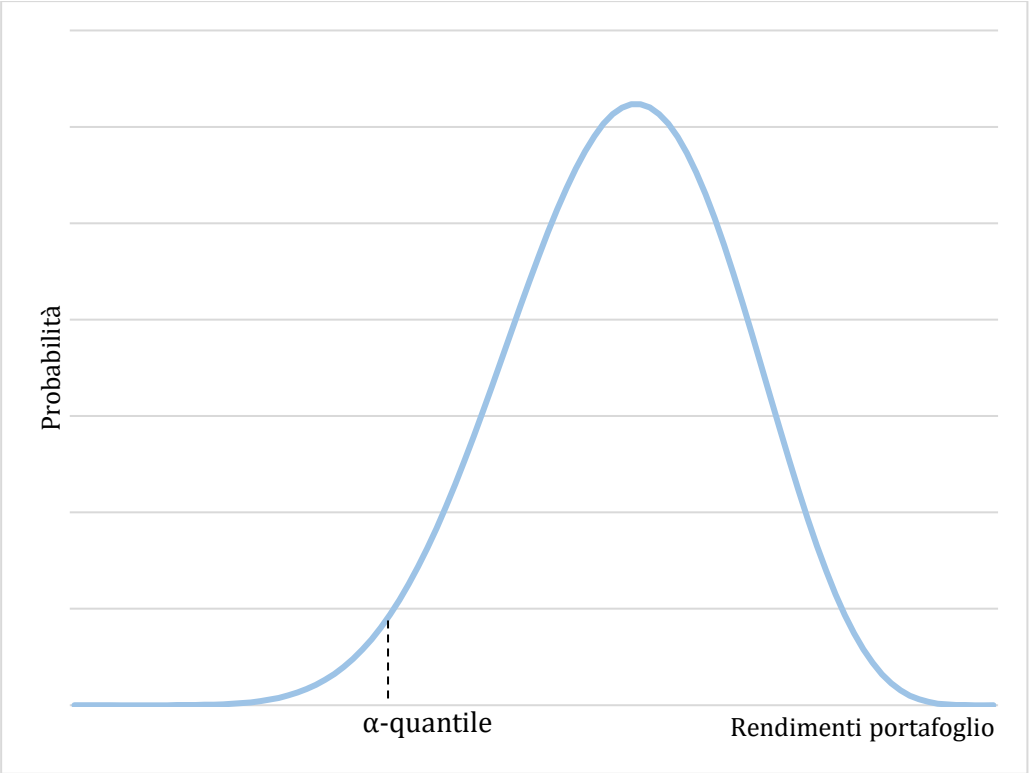


Figura 8: distribuzione dei rendimenti del portafoglio B.

Come si può osservare, il portafoglio A presenta una distribuzione asimmetrica a destra, dove le perdite sono limitate mentre i potenziali guadagni sono illimitati. All'opposto, il secondo portafoglio B presenta delle perdite illimitate a discapito di guadagni limitati. Nonostante sia piuttosto evidente che qualsiasi investitore razionale considererebbe il secondo portafoglio più rischioso rispetto al primo, l'utilizzo del VaR come misura di rischio porta a trarre la conclusione errata che i due portafogli siano rischiosi allo stesso modo.

Un altro problema del VaR deriva dal suo comportamento nel momento in cui vengono aggregate più fonti di rischio, e la loro distribuzione congiunta non è gaussiana. In particolare, diversamente dalla varianza, esso non rispetta l'assioma di subadditività, in quanto non riesce sempre ad individuare la contrazione di rischio che si genera formando portafogli con titoli poco correlati. Ad esempio, il VaR ottenuto da un portafoglio formato da due titoli con distribuzione non normale potrebbe essere maggiore della somma del VaR dei due singoli titoli. Pertanto, la non-subadditività ostacola la diversificazione del portafoglio.

Sebbene questa misura di rischio sia stata introdotta nel tentativo di misurare la rischiosità associata a "particolari" distribuzioni, paradossalmente dai suddetti limiti ne deriva che può essere utilizzata in modo corretto solamente in presenza di

distribuzioni normali, ovvero nella stessa condizione in cui può essere impiegata anche la varianza (di più semplice calcolo) come misura di rischio (Szegö, 2002).
Infine, da un punto di vista computazionale, non è agevole utilizzare il VaR nei problemi di ottimizzazione vincolata: il problema di programmazione stocastica che ne risulta è piuttosto difficile da risolvere. Inoltre diversi studi[16] hanno mostrato come l'utilizzo di modelli *VaR-based*, con l'obbiettivo di massimizzare l'utilità degli investitori, formano portafogli che non sono adeguatamente immuni agli shock di mercato.
A causa dei limiti descritti in questo paragrafo, nel presente lavoro si è deciso di non utilizzare il VaR come misura di rischio nella formulazione del modello di ottimizzazione vincolata, che verrà esposto nel seguito. La misura di rischio scelta ricade all'interno dell'insieme di misure che rispettano congiuntamente i quattro assiomi di coerenza.

2.2 Gli assiomi di coerenza

Formulare degli assiomi significa fissare, all'interno di un insieme di precise affermazioni, l'intrinseca natura di un concetto. È un passo che deve essere intrapreso nel processo di traduzione di una realtà complessa in una formulazione matematica (Acerbi e Tasche, 2002).
Gli assiomi di coerenza, introdotti da Artzner, Delbaen, Eber e Heath nel 1998, sono quattro proprietà fondamentali, che devono essere soddisfatte da una generica misura di rischio per essere definita coerente: l'inosservanza anche di solo uno di tali assiomi porta sempre a risultati paradossali e sbagliati.
Prima di esporli in modo dettagliato, si rende tuttavia necessario fornire una definizione generale di *misura di rischio*:

- misurare il rischio equivale a stabilire una corrispondenza ρ tra lo spazio X delle variabili casuali e un numero reale non negativo, $\rho : X \to R$ (Szegö, 2002).

Di conseguenza affinché la funzione ρ sia definibile come misura di rischio deve

[16] Basak e Shapiro, 2001, Yamai e Yoshiba, 2002.

rispettare tre proprietà di base:

- *positività*, in quanto una misura di rischio deve sempre associare ad una variabile casuale un valore strettamente positivo, oppure nullo nel caso in cui non ci sia alcuna forma di rischio (mai valori negativi poiché non avrebbero senso);
- *linearità*, ovvero la misura di rischio deve essere connessa in maniera lineare con il rendimento;
- *convessità*, ossia la relazione

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

dove R_x e R_y sono due parametri aleatori e $\vartheta \in [0{:}1]$.

Quest'ultima proprietà afferma che il rischio di un portafoglio composto da due titoli deve sempre essere non maggiore della somma pesata dei singoli rischi associati ai due titoli, con pesi equivalenti alle percentuali d'investimento. La mancanza di convessità comporterebbe la non convenienza a sviluppare portafogli composti da più titoli.

A questo punto si possono definire le quattro proprietà aggiuntive e complementari, rispetto a quelle basilari appena esposte, la cui osservanza garantisce sia un'esatta comprensione del concetto di rischio connesso ad un'attività finanziaria, sia la coerenza della misura di rischio stessa:

1. *Invarianza transizionale*: considerando inizialmente un portafoglio composto solo da attività rischiose, il suo rischio complessivo può essere contratto di un valore pari ad α investendo lo stesso valore in un titolo privo di rischio:

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

Da questo ne deriva che costruendo un portafoglio composto da un titolo rischioso e da un titolo *risk-less* di valore uguale al rischio del primo titolo, si ottiene che il rischio complessivo di tale portafoglio è zero:

$$\rho\big(X + \rho(X)\big) = 0.$$

In altri termini, nell'ambito delle misure di rischio coerenti, la presenza di un titolo a rendimento certo consente di abbassare il rischio.

La varianza è un esempio di misura di rischio che non soddisfa tale assioma, in quanto la varianza della somma tra un titolo rischioso e un titolo privo di rischio (una costante) è pari alla varianza del titolo rischioso.

2. *Subadditività*: afferma che il rischio di un portafoglio composto da due titoli X e Y deve essere sempre minore o uguale della somma del rischio dei due singoli titoli:

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

Pertanto se c'è la possibilità di contrarre il rischio, la misura che si sta utilizzando deve essere in grado di coglierla[17].

Un risultato implicito, ricavabile da questo assioma, è che la diversificazione di portafoglio porta ad una contrazione del rischio solamente se la misura di rischio utilizzata rispetta tale assunto. Come evidenziato nel paragrafo precedente, il VaR non sempre ottempera questa proprietà quando le distribuzioni oggetto di analisi non seguono un andamento normale.

3. *Positiva omogeneità*: garantisce che il rischio segua l'investimento effettuato. Dato un numero non negativo λ, il rischio di λ volte un investimento deve essere uguale a λ volte il rischio dell'investimento stesso:

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

In altri termini, aumentando l'investimento in un titolo a rendimento aleatorio, aumenta in modo proporzionale anche il rischio connesso a quell'investimento.

4. *Monotonicità*: può essere definita come una via di mezzo fra la qualità di un investimento e il "sentire" dell'investitore. Se un individuo preferisce l'investimento X a Y per un qualche motivo, allora la misura di rischio deve essere in grado di recepire la preferenza di quell'individuo. Più formalmente, date due variabili aleatorie X e Y, con $Y \preccurlyeq X$ (X preferita a Y), allora il rischio associato alla variabile X sarà minore del rischio associato alla variabile Y:

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

Più dettagliatamente, qualora esistessero opportunità di arbitraggio[18] sui mercati finanziari, la misura di rischio deve essere in grado di sfruttarli

[17] Come si può notare è molto simile alla proprietà di *convessità*.
[18] Con il termine arbitraggio si intende la capacità di realizzare un profitto certo senza che l'investitore, che realizza tale operazione, vada incontro ad alcun rischio.

attribuendo al titolo meno redditizio un maggior rischio.

Dato che la proprietà di *convessità* deriva dagli assiomi di *positiva omogeneità* e *subadditività*, quest'ultimi si rendono fondamentali anche da un punto di vista di ottimizzazione di portafoglio. Difatti, affinché sia garantita un'unica soluzione del problema di minimizzazione, è essenziale che la misura di rischio sia convessa.
I quattro assiomi di coerenza non si limitano ad individuare un'unica misura di rischio coerente, ma circoscrivono l'insieme di un ampio gruppo di misure. Data la loro importanza operativa, l'obbiettivo del prossimo paragrafo sarà proprio quello di presentare le principali misure di rischio coerenti sviluppate in letteratura.

2.3 Alcune misure di rischio coerenti

2.3.1 Tail Conditional Expectation e Worst Conditional Expectation

In letteratura sono state proposte una serie di misure di rischio che rispettano tutti gli assiomi di coerenza. Le prime due, che seguono queste linee guida, furono proposte dagli stessi Artzner *et Al.* nel lavoro dove introdussero le quattro proprietà. La prima, detta Tail Conditional Expectation (TCE), corrisponde al valor medio di tutte le realizzazioni che sono minori o uguali al VaR:

$$TCE_a(X) \stackrel{\text{def}}{=} -E[X|X \leq -VaR_a(X)];$$

il segno meno ha il solo scopo di rendere positiva la misura, in quanto, normalmente, si lavora con le perdite. Nonostante sia una misura preferibile al VaR, essa non rispetta l'assioma di subadditività quando le distribuzioni di probabilità in esame non sono continue. Lo stimatore naturale del TCE è dato da:

$$TCE_n^{(a)}(X) = -\frac{\sum_{i=1}^{n} X_i \mathbf{1}_{\{X_i \leq -VaR_a(X)\}}}{\sum_{i=1}^{n} \mathbf{1}_{\{X_i \leq -VaR_a(X)\}}},$$

dove $\mathbf{1} = \begin{cases} 1 \ se\ la\ relazione \text{ è } vera \\ 0\ se\ la\ relazione \text{ è } falsa \end{cases}$.

La seconda misura, detta Worst Conditional Expectation (WCE), calcola il valore medio delle perdite dato un evento A, dove la probabilità di A è maggiore di α:

$$WCE_a(X) \stackrel{\text{def}}{=} -inf\{E[X|A]|P[A] > \alpha\}.$$

Quindi non si limita ad osservare la distribuzione, ma considera tutti i possibili eventi A che siano coinvolti in perdite: tutti questi eventi devono avere una probabilità maggiore di α. In altre parole, seleziona la perdita più elevata che si potrebbe generare qualora si verificasse un evento probabile A.
L'apparente similitudine di queste due misure di rischio non deve trarre in inganno. Se da un lato il WCE è preferibile al TCE in quanto rispetta sempre tutti gli assiomi di coerenza, dall'altro lato il primo non è molto adottato nella pratica poiché il suo utilizzo necessita la conoscenza profonda dello spazio di probabilità degli eventi.
Queste due misure di rischio si focalizzano entrambe sulla forma della coda sinistra della distribuzione dei rendimenti e ne prendono il valor medio. Inoltre, Artzner *et Al.* dimostrano la validità della seguente disuguaglianza:

$$TCE_a \leq WCE_a.$$

2.3.2 Expected shortfall e Conditional Value at Risk

L'obiettivo della formulazione di una misura di rischio caratterizzata dalla peculiarità di coniugare contemporaneamente i pregi delle due precedenti misure fu raggiunto mediante l'introduzione dell'Expected Shortfall (ES).
Individuando con $F(X)$ la funzione di densità di probabilità della variabile casuale X tale che $P(X \leq x)$, ed introducendo la funzione inversa di $F(X)$

$$F^{-1}(a) = inf\{x\ |F(x) \geq \alpha\},$$

è dimostrato[19] che l'ES può essere espresso come

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

Quindi è il valor medio della coda sinistra della distribuzione oltre un certo valore soglia definito dal VaR. In altri termini l'ES afferma che considerando l'ultimo $1 - \alpha\%$ della coda, il suo valore è pari alla media delle perdite all'interno di quell'area.
La definizione originale, ripresa da Acerbi e Tasche, nonostante sia molto più rigorosa rispetto alla precedente, è uguale nella sostanza: data una distribuzione

[19] Si veda Acerbi e Tasche, 2002.

delle *profit and loss X*, e specificato un determinato *holding period*, l'ES con livello di significatività $\alpha \in [0:1]$ è definito come:

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

dove $x^{(\alpha)}$ è pari al valore del VaR_α.

Il secondo addendo è un correttivo da sottrarre al valor medio (ossia il primo addendo) nel momento in cui la probabilità di $X \leq x^{(\alpha)}$ è maggiore di $1-\alpha$: tale eventualità si manifesta, in particolare, nel caso di distribuzioni discrete. In presenza di distribuzioni di probabilità continue, ovvero quando $P\left(X \leq x^{(\alpha)}\right) = 1-\alpha$, si può dimostrare che l'ES coincide con il TCE.

Una rappresentazione grafica dell'ES è mostrata in Figura 9. Come si può notare tale misura si spinge oltre la definizione di una soglia, offrendo una valutazione del rischio più corretta rispetto al VaR. A riprova di ciò, diversi studi, come quello di Liang e Park (2007), hanno fornito un'evidenza empirica sulla superiorità dell'ES rispetto al VaR come misura del *downside risk.*

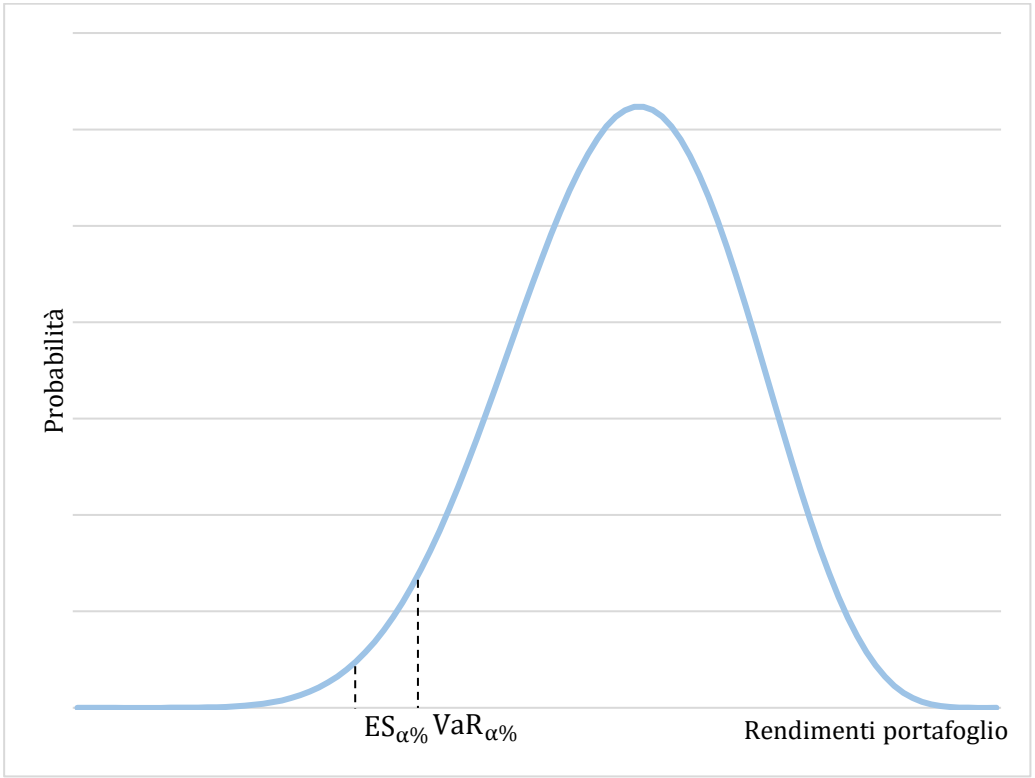


Figura 9: ES e VaR di una distribuzione di probabilità dei rendimenti di un portafoglio.

L'ES è senza dubbio la misura di rischio coerente più apprezzata ed utilizzata tra

quelle proposte in letteratura. Da un punto di vista di ottimizzazione di portafoglio, un ulteriore vantaggio dell'ES deriva dalla sua semplicità di applicazione: il problema di minimizzazione di questa misura di rischio è convesso.
Per quanto riguarda la sua stima si possono utilizzare i metodi parametrici o quelli non parametrici. Il metodo più utilizzato a livello operativo appartiene al secondo gruppo ed è la simulazione storica: partendo dalla serie storica delle performance del portafoglio che si sta analizzando e, dopo averle ordinate in maniera crescente o decrescente, si prende l'$1-\alpha\%$ delle maggiori perdite (in termini negativi) e si calcola il valor medio di quest'ultime:

$$ES_\alpha(X) = -\frac{\sum_{i=1}^{W} x_{i:n}}{W},$$

dove W è la parte intera di $n(1-\alpha\%)$.

Questo rappresenta lo stimatore campionario dell'ES, ed è *BLUE*[20] solo in presenza di distribuzioni normali.
All'interno del gruppo dei metodi parametrici si possono trovare differenti criteri per il calcolo dell'ES, che variano in base alla distribuzione di probabilità prevista per i profitti e le perdite di un determinato titolo o portafoglio. Ipotizzando, a titolo esemplificativo, che la distribuzione sia una *t di Student*, allora l'ES può essere espresso come (Broda e Paolella 2011):

$$ES_p(X) = \frac{1}{p} Ttail(q_p, n),$$

dove:

- $Ttail(c,n) = -\frac{n^{-1/2}}{\sigma B\left(\frac{n}{2},\frac{1}{2}\right)}\left(1+\frac{c^2}{n}\right)^{-\frac{n+1}{2}}\frac{n+c^2}{n-1}$;
- $c > 0$ è il parametro di scala;
- $n > 0$ sono i gradi di libertà.

Nonostante il superamento dei limiti delle precedenti misure di rischio, e gli indiscutibili benefici derivanti dal suo impiego, anche l'ES presenta dei problemi. Ad esempio, Yamai e Yoshiba (2002) evidenziano come l'ES necessiti di un grande campione per mantenere gli stessi livelli di accuratezza del VaR. Inoltre, non

[20] Best Linear Unbiased Estimator – Il migliore stimatore lineare e non distorto.

esamina i momenti di ordine superiore che consentirebbero di comprendere in modo completo le distribuzioni non-normali.
Il Conditional Value-at-risk (CVaR), proposto da Rockafellar e Uryasev (2000), è una formulazione alternativa all'ES. Considerando variabili casuali continue, può essere definito come il valore atteso delle perdite che eccedono il VaR:

$$CVaR_k = \phi_k(x) = (1-k)^{-1} \int_{f(x,y)\geq\alpha_k(x)} f(x,y)p(y)dy,$$

dove $k \in [0:1]$ è il livello di significatività scelto e $\alpha_k(x)$ rappresenta il VaR. Un ulteriore formulazione è

$$CVaR_k = VaR_k + E[f(x,y) - VaR_k | f(x,y) > VaR_k].$$

Come si può notare queste definizioni non offrono alcun suggerimento su come calcolare il CVaR senza conoscere a priori il VaR. Rockafellar e Uryasev riuscirono ad oltrepassare tale limite adottando la seguente formula:

$$F_k(x,\alpha) = \alpha + (1-k)^{-1} \int_{y\in R^m} [f(x,y) - \alpha]^+ p(y)dy,$$

dove $[u]^+ = u$ se $u > 0$ e $[u]^+ = 0$ se $u \leq 0$. Oltre a dimostrare che $F_k(x,\alpha)$ è convessa e differenziabile, provarono che il CVaR può essere calcolato come:

$$\phi_k(x) = \min_{\alpha\in R} F_k(x,\alpha);$$

inoltre $\min_{x\in X} \phi_k(x) = \min_{(x,\alpha)\in X\times R} F_k(x,\alpha)$.

Questo è un risultato di considerevole importanza poiché consente di determinare il CVaR di un portafoglio anche senza dover calcolare il relativo VaR.

2.4 Misure di rischio two-sided

L'ultima classe di misure di rischio coerenti, che verranno trattate nel presente lavoro, sono le cosiddette *Two-sided risk measures* introdotte da Chen e Wang nel 2008.
Ciò che contraddistingue tali misure dalle precedenti è la loro capacità di combinare in modo differente sia gli scostamenti positivi che quelli negativi rispetto al valor

medio. In altri termini, da un lato, mirano a gestire il rischio "negativo", ovvero le performance al di sotto della media, dall'altro, cercano di sfruttare la coda "buona" della distribuzione. Tale definizione sposa l'ottica fondamentale della teoria dei giochi[21]: nella valutazione del rischio di un determinato titolo si rende necessario considerare entrambi i comportamenti sia del venditore che dell'acquirente.

Prima della loro introduzione, le uniche misure di rischio "Two-sided" proposte in letteratura erano quelle in Hamza e Janssen (1998), dove il rischio è misurato attraverso la combinazione convessa delle semivarianze dei rendimenti del portafoglio. Oltre ad essere non coerenti, tali misure presentano il limite di essere simmetriche poiché si utilizzano le semivarianze per misurare sia le deviazioni positive che negative dalla media.

In termini analitici la misura di rischio *Two-sided* proposta da Chen e Wang può essere definita come:

$$\rho_{\alpha,p}(X) \stackrel{\text{def}}{=} \alpha\sigma_1^+(X) + (1-\alpha)\sigma_p^-(X) - E[X]$$
$$= \alpha\|(X - E[X])^+\|_1 + (1-\alpha)\|(X - E[X])^-\|_p - E[X]$$

dove:

- $\|X\|_p \coloneqq (\sum_{i=1}^{n}|x_i|^p)^{\frac{1}{p}}$;
- $p \in [1, \infty)$;
- $0 \leq a \leq 1$.

I primi due addendi possono essere definiti come delle distanze rispetto alla media delle varie realizzazioni, calcolate con esponenti diversi. In particolare, il primo addendo è fissato di ordine uno, quindi tale norma è il valore assoluto e, con appropriati aggiustamenti, può essere definita come il MAD di tutte le realizzazioni sopra la media. La scelta di elevarlo a 1 deriva dal fatto che si stanno valutando *sovra-performance*, di conseguenza non si vuole creare un effetto distorsivo utilizzando ordini di livello più elevato. Il secondo addendo, all'opposto, considera la distanza di tutte le realizzazioni inferiori alla media. In questo caso si utilizza un generico ordine *p* poiché, mentre con i guadagni è preferibile essere obiettivi, per le perdite si vuole catturare (entro certi limiti) l'atteggiamento del singolo individuo

[21] La teoria dei giochi è una disciplina che studia le decisioni di un individuo in situazioni di iterazione o conflitto con altri soggetti rivali.

nei confronti del rischio. Infine l'ultimo addendo garantisce la coerenza di tale misura.

I due parametri α e p, dai quali dipende la formula, hanno lo scopo di esprimere in modo flessibile le differenti attitudini al rischio degli investitori: il primo è una costante che quantifica la preferenza dell'investitore tra la volatilità negativa e la volatilità positiva; il secondo è un fattore locale che, oltre a riassumere l'avversione al rischio del singolo investitore, include le informazioni connesse con le caratteristiche della distribuzione dei rendimenti (come la curtosi o l'asimmetria).

Più dettagliatamente, tale misura di rischio può essere espressa come una combinazione lineare pesata di momenti positivi e momenti negativi della distribuzione di probabilità dei rendimenti, dove i pesi sono pari ad α e $1-\alpha$.

Chen e Wang dimostrano i seguenti teoremi:

- Per ogni $p_i \geq 1$, $0 \leq \alpha_i \leq 1$, $\lambda_i \in (0,1)$, $i = 1, \dots, \infty$, con $\sum_{i=1}^{\infty} \lambda_i = 1$ la misura di rischio seguente è coerente

  $$\rho_{a,p}(X) = \sum_{i=1}^{\infty} \lambda_i \, \rho_{\alpha_i p_i} = -E[X] + \sum_{i=1}^{\infty} \lambda_i \, \alpha_i \sigma_1^+ + \sum_{i=1}^{\infty} \lambda_i \, (1-\alpha_i)\sigma_{p_i}^-,$$

  dove $q = sup\{p_i | p_i \geq 1, i = 1, \dots, \infty\}$. Questo significa che può essere utilizzata anche la combinazione convessa "generalizzata" di molteplici $\rho_{\alpha,p}$ come misura di rischio, sfruttando in questo modo i vantaggi derivanti dall'impiego simultaneo di differenti $\rho_{\alpha,p}$. Ad esempio, se si effettua la combinazione convessa di $\rho_{\alpha_1,2}$ $(0 \leq \alpha_1 \leq 1)$ e $\rho_{\alpha_2,\infty}$ $(0 \leq \alpha_2 \leq 1)$, si ottiene una nuova misura di rischio che possiede le proprietà dell'*absolute deviation*, della semi-varianza e delle *maximum loss risk measures*.
- La misura di rischio coerente $\rho_{\alpha,p}$ non decresce all'aumentare del fattore p, e non cresce all'aumentare del fattore α. Pertanto un investitore che adotta questa misura di rischio dovrebbe considerare X più rischioso rispetto ad un altro investitore che utilizza a sua volta $\rho_{\alpha,p}$ ma con un livello di p inferiore. Dall'altro lato una minore avversione al rischio dell'investitore dovrebbe essere modellizzata da un maggior valore di α (a parità di p).
- Le misure di rischio $\rho_{\alpha,p}$ e MAD rispettano la relazione

  $$\rho_{\alpha,p}(X) + E[X] \geq \frac{MAD(X)}{2},$$

dove si verifica l'uguaglianza se e solo se $p = 1$. Diversamente la seguente disuguaglianza definisce la relazione tra $\rho_{\alpha,p}$ e il $CVaR$

$$\rho_{\alpha,1}(X) < CVaR_{\gamma}(X),$$

dove $0 \leq \gamma \leq \alpha < 1$, con γ pari al livello di confidenza. Inoltre,

$$CVaR_{\alpha}(X) - \rho_{\alpha,1}(X) = \left(\frac{1}{\alpha} - 1\right) E[(X - E[X])^{-}].$$

Da un punto di vista di ottimizzazione di portafoglio la misura in oggetto non richiede notevoli sforzi computazionali.

Nel modello di ottimizzazione formulato nel presente lavoro si utilizzerà una misura di rischio two-sided. Nel successivo capitolo si introdurranno le due metaeuristiche per la risoluzione del problema di allocazione.

CAPITOLO 3

PARTICLE SWARM OPTIMIZATION E BACTERIAL FORAGING OPTIMIZATION

In questo capitolo si introdurranno le due metaeuristiche che verranno utilizzate nel seguito per la risoluzione del problema di ottimizzazione vincolata proposto da DeltaHedge: la Particle Swarm Optimization (PSO) e la Bacterial Foraging Optimization (BFO).

3.1 Euristiche ed algoritmi esatti

L'obiettivo di questo elaborato è quello di determinare la soluzione ottima a un particolare problema di selezione di portafoglio presentato dalla società *DeltaHedge*[22]. In genere, per la risoluzione di problemi di questo tipo, si possono seguire due strade non necessariamente alternative: i metodi esatti (o completi) e i metodi euristici.

Il primo gruppo incorpora tutti quei metodi che risolvono uno specifico problema di ottimizzazione determinando la soluzione ottima tra tutte quelle ammissibili. L'utilizzo di tali algoritmi[23], nonostante permetta di ottenere soluzioni precise e puntuali, non oltrepassa il confine delimitato dalla risoluzione del particolare problema per cui sono sviluppati. La loro formulazione richiede solide competenze matematiche che spesso i dipendenti delle società che dovrebbero implementare tali modelli non possiedono. Inoltre l'impiego di questi metodi è compromesso in presenza di problemi la cui complessità è di tipo NP-Hard.[24]

I limiti appena descritti favoriscono il ricorso al secondo gruppo di metodi, ovvero gli algoritmi euristici. Quest'ultimi consentono di ottenere una soluzione "accettabile" del problema in tempi computazionali relativamente brevi. Perciò, se

[22] Vedi capitolo 4.
[23] Un algoritmo è un procedimento di risoluzione di un problema in una limitata quantità di tempo, mediante un insieme finito di operazioni non ambigue ed effettivamente calcolabili.
[24] Indentifica problemi che sono complessi almeno quanto i problemi *NP-Completi*. Per tali problemi non sono noti algoritmi di complessità polinomiale che siano in grado di risolverli.

da un lato non garantiscono la coincidenza tra la soluzione trovata e quella ottima, dall'altro lato presentano una maggiore flessibilità (rispetto ai metodi precedenti), permettono di risolvere problemi di tipo NP-Hard e, generalmente, ottengono buone soluzioni sub-ottime.

Per molti problemi di ottimizzazione si possono sviluppare dei metodi euristici specifici che mirano a sfruttare le proprietà del singolo problema da risolvere. Tuttavia, negli ultimi anni, l'attenzione (a livello sia accademico che applicativo) si è rivolta verso i metodi euristici di tipo generale, i quali forniscono dei risultati che nella maggior parte dei casi sono migliori rispetto a quelli ottenuti dalle *euristiche specifiche*.

In letteratura si possono trovare molte tecniche di questo tipo, al punto che non esiste una sistemazione e classificazione universalmente accettata. Un possibile ordinamento, sicuramente incompleto, è il seguente:

- Euristiche costruttive: partendo da un insieme delle soluzioni vuoto, consentono di costruire la soluzione in modo progressivo, ovvero aggiungendo all'insieme un singolo elemento ad ogni iterazione. La loro applicazione ideale ricade sulla ricerca di quelle soluzioni che possono essere ottenute come sottoinsieme di più elementi. Lo schema concettuale di tali algoritmi è il seguente:
  1. Scelta di un input di partenza (fase di inizializzazione) per formare una soluzione parziale.
  2. Selezione di un secondo input, mediante un criterio predeterminato, che deve essere aggregato alla soluzione parziale precedente.
  3. Nel caso di non ammissibilità della soluzione si torna al punto precedente, altrimenti si arresta l'algoritmo.

  Esempi di euristiche di questo tipo sono gli algoritmi *greedy* e gli algoritmi che sfruttano le tecniche di ottimizzazione esatta.
- Euristiche migliorative (o di ricerca locale): partendo da una soluzione iniziale, tali algoritmi cercano di migliorarla iterativamente perlustrando l'intorno (che deve essere definito) della soluzione stessa. Lo schema concettuale delle euristiche migliorative è il seguente:
  1. Scelta di una soluzione iniziale per avviare il processo di ricerca.

2. Determinazione dell'intorno della soluzione iniziale mediante un metodo specifico. In particolare, un intorno (o vicinato) è un'applicazione $N: s \rightarrow N(s)$ che ad ogni punto s della regione ammissibile accosta un sottoinsieme $N(s)$ della regione ammissibile X.
3. Valutare tutte le soluzioni dell'intorno mediante un'apposita funzione e scegliere quella migliore. La soluzione minima locale è quella per cui il valore della sua funzione obiettivo è inferiore rispetto allo stesso delle altre funzioni obiettivo associate alle rimanenti soluzioni del vicinato.
4. Arresto del processo se la soluzione trovata è peggiore di quella iniziale, altrimenti si torna al secondo passo sostituendo la soluzione iniziale con quella migliore.

Nella maggior parte dei casi la soluzione trovata mediante questi algoritmi rappresenta solamente un *ottimo locale* e non *globale*.

- Iper-euristiche: si tratta di algoritmi che individuano metodi di ottimizzazione con il fine di adattarli *automaticamente* a più problemi. Sono temi la cui ricerca è ancora in una fase iniziale.

Si rende necessario evidenziare che la preferibilità di un metodo euristico, rispetto ad uno esatto, dovrebbe sempre essere il risultato di una preventiva analisi basata sulle caratteristiche e dimensioni del problema.
Nel prossimo paragrafo verrà presentata ed approfondita un'ulteriore classe dei metodi euristici, che può essere definita come uno sviluppo delle tecniche di ricerca locale.

3.2 Metaeuristiche

Le metaeuristiche (dal greco *meta*, che significa "oltre", e *heuriskein* che significa "cercare") sono una nuova tipologia di algoritmi approssimati che si sono diffusi, in particolare, negli ultimi 30 anni. Questi mirano a combinare diversi metodi euristici elementari in strutture più complesse, che garantiscono un'esplorazione efficace ed efficiente dello spazio di ricerca.
In letteratura sono state proposte molte definizioni di metaeuristica, senza, tuttavia, convergere ad una definizione che sia universalmente condivisa. Blum e Roli (2003)

riassumono le caratteristiche fondamentali di tali algoritmi nelle seguenti proprietà:

- Sono strategie che guidano il processo di ricerca;
- L'obiettivo è quello di esplorare in modo efficiente lo spazio di ricerca per trovare soluzioni ottime;
- Sono costituite da algoritmi approssimati e (di solito) non deterministici;
- Incorporano metodi che evitano di rimanere intrappolati in determinate aree dello spazio di ricerca;
- Sono costituite da tecniche che spaziano dalle semplici procedure di ricerca locale ai più complessi processi di apprendimento;
- Le metaeuristiche più avanzate sfruttano la *memoria* per guidare la loro ricerca.

In altri termini, questi algoritmi non si limitano ad avvicinarsi alla soluzione ottima (come le euristiche precedenti), ma cercano di esplorare più approfonditamente quelle zone che risultano più interessanti all'interno dello spazio delle soluzioni. Tali concetti sono riassunti nelle due caratteristiche che contraddistinguono le metaeuristiche, ovvero l'intensificazione e la diversificazione. Con il primo termine si intende l'attitudine dell'algoritmo ad approfondire la ricerca nelle aree più interessanti all'interno dello spazio delle soluzioni; diversamente, il secondo termine, si riferisce alla capacità di individuare zone non visitate fino a quel momento, con il fine di scoprire nuove aree attraenti dove incrementare la ricerca. Il corretto bilanciamento di questi due aspetti consente sia di indirizzare la ricerca verso l'identificazione di una moltitudine di minimi locali, sia di evitare di sprecare tempo in zone poco promettenti o già esplorate.
Prima di definire una possibile classificazione delle metaeuristiche, si rende necessario evidenziare che negli ultimi anni si è assistito ad una eccessiva proliferazione delle stesse.
In molti casi queste differiscono solo per le suggestive metafore su cui sono fondate, ma non sulla logica sottostante ai meccanismi che consentono di raggiungere buone soluzioni: *«an algorithm is good if it provides good results, and not if it is described by a suggestive metaphor»* (Sörensen, 2015). Le metaeuristiche possono essere classificate nel seguente modo:

- *Single point search vs. population-based*: differiscono per il numero di soluzioni che vengono sfruttate nello stesso momento. In particolare, nel primo caso viene generata una singola soluzione che si muove all'interno dello spazio di ricerca, mentre, nel secondo caso, il processo di ricerca è eseguito mediante la generazione e gestione di un insieme di soluzioni detto "popolazione".
- *Nature-inspired vs. non-nature inspired*: tale distinzione si fonda sulla metafora a cui è ispirata la metaeuristica. Algoritmi che si basano sulla natura sono, ad esempio, gli *algoritmi genetici* e gli *Ant Algorithms*, che si ispirano rispettivamente all'evoluzione biologica e alle colonie di formiche. All'opposto, esempi di algoritmi che non prendono spunto dalla "natura" sono i *Fireworks Algorithms* e la *Tabu Search*. In letteratura si possono trovare metodi ibridi a cui non è possibile dare una specifica classificazione all'interno di questa distinzione.
- *Dynamic vs. static objective function*: differiscono per il modo in cui viene utilizzata la funzione obiettivo. Nel primo caso rimane sempre la stessa per tutto il processo di ricerca, mentre, nel secondo caso, questa si modifica con l'obbiettivo di incorporare le informazioni raccolte durante il processo fino a quel momento.
- *Memory usage vs. memory-less methods*: distinzione basata sull'utilizzo o meno della storia della ricerca da parte delle metaeuristiche. Gli algoritmi appartenenti al primo gruppo possono essere ulteriormente suddivisi in algoritmi con memoria di breve periodo e algoritmi con memoria di lungo periodo. I primi tengono traccia delle soluzioni visitate più recentemente, mentre i secondi solitamente sono un accumulo di parametri di sintesi sulla ricerca. Gli algoritmi appartenenti al gruppo *memory-less* utilizzano solamente il corrente stato del processo di ricerca per determinare l'azione successiva (processo di Markov).
- *One vs. various neighborhood structures*: distinzione che riguarda la topologia dello spazio di ricerca. Mentre per alcuni algoritmi la struttura del vicinato non cambia durante il processo, per altre metaeuristiche questa si modifica dinamicamente con l'obiettivo di diversificare il più possibile la ricerca.

Prima di presentare in modo dettagliato le due metaeuristiche che verranno utilizzate nel presente lavoro per risolvere il problema di ottimizzazione vincolata, si procederà con una breve descrizione delle due più importanti classi di algoritmi appartenenti all'ampia famiglia delle metaeuristiche.

3.2.1 Gli algoritmi evolutivi

Nonostante nella letteratura si possano trovare differenti varianti degli algoritmi evolutivi (EA), nel presente paragrafo, per semplicità, ci si limiterà a definire la logica comune che caratterizza tali algoritmi.

Gli EA mirano a risolvere problemi di ottimizzazione complessi ispirandosi alla capacità della natura di far evolvere gli esseri viventi adattandoli all'ambiente in cui si trovano. I due principi di selezione naturale e di sopravvivenza dei soggetti più adatti, sui quali si fondano gli EA, conducono all'inevitabile collegamento di quest'ultimi alla teoria dell'evoluzione della specie formulata dal geologo, naturalista ed illustratore britannico Charles Darwin (1859).

Il funzionamento di tali algoritmi parte dalla generazione casuale di una popolazione di individui, i quali vengono valutati sulla base della loro fitness (questa può essere il valore di una funzione obiettivo o di qualsiasi altro tipo di misura di qualità). Dopo aver selezionato gli individui migliori, il passo seguente consiste nel creare la popolazione della successiva iterazione attraverso due metodi alternativi: la ricombinazione, ovvero mettere insieme due o più individui con il fine di produrre nuovi soggetti; la mutazione (o modificazione), che determina un auto-adattamento degli individui della precedente generazione. Tale processo può essere replicato finché non si ottiene un livello soddisfacente della funzione di fitness, ovvero fino al momento in cui un predeterminato criterio di arresto non viene raggiunto.

Gli individui, che formano la popolazione ad ogni iterazione, non sono necessariamente delle soluzioni al problema che si intende risolvere. Difatti questi possono anche essere delle soluzioni parziali, oppure qualsiasi altro tipo di oggetto che può essere trasformato in una o più soluzioni in un modo strutturato. La Figura 10 mostra la struttura di base per tutti gli EA.

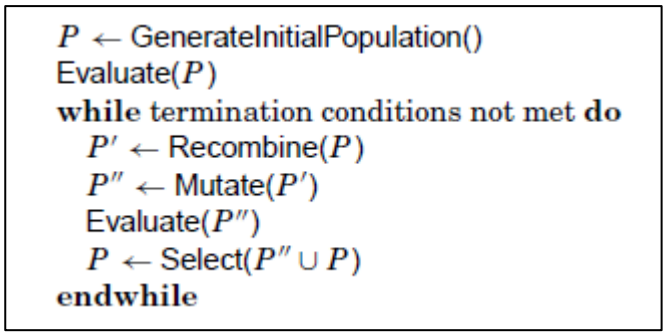


Figura 10: codice generico EA. Fonte: Blum and Roli (2003).

Individui con un miglior valore della fitness hanno una maggiore probabilità di essere scelti come membri della popolazione nella successiva generazione (o come genitori della nuova generazione di individui). Questo corrisponde al principio di *survival of the fittest* nell'evoluzione naturale.

I principali sottogruppi di algoritmi appartenenti alla famiglia degli EA sono i seguenti:

- Evolution Strategies (ES);
- Learning Classifier System (LCS);
- Genetic Algorithms (GE);
- Evolutionary Programming (EP);
- Programmazione genetica (GP).

Un'ulteriore classe di algoritmi, strettamente connessa con la famiglia degli EA, senza farne effettivamente parte, è la Swarm Intelligence.

3.2.2 Swarm Intelligence

La Swarm Intelligence (SI) raggruppa un insieme di algoritmi ispirati dai comportamenti collettivi messi in atto da alcuni gruppi di insetti o animali. I primi a proporre tale espressione furono Hackwood, Wang e Beni (1988) nell'ambito dei sistemi robotici cellulari.

In particolare le tecniche di SI hanno lo scopo di determinare la soluzione ottima di un determinato problema, sfruttando il comportamento globale di uno "sciame" di agenti omogenei e non sofisticati. Quest'ultimi condividono informazioni ed esperienze l'uno con l'altro, riuscendo, in questo modo, a risolvere compiti anche molto complessi. Il concetto chiave che caratterizza tali sciami è l'auto-

organizzazione, ovvero un sistema che nel corso del tempo punta a perfezionare le sue capacità migliorando l'organizzazione dei suoi elementi. Un ulteriore aspetto fondamentale è la decentralizzazione, ossia l'assenza di un individuo che ha il compito di supervisionare il funzionamento dell'intero processo.

Si può affermare che "l'intelligenza dello sciame" deriva da iterazioni tra semplici agenti che, se presi individualmente, non sarebbero in grado di risolvere nemmeno i compiti più banali. Il comportamento delle formiche è un chiaro esempio di quanto definito: nonostante la loro intelligenza molto limitata, queste possono mostrare straordinarie ed inaspettate caratteristiche a livello di gruppo mediante una particolare forma di comunicazione implicita (detta *stigmergia*). Ulteriori esempi di SI che si possono incontrare in natura sono i comportamenti delle colonie di api, degli stormi di uccelli, delle popolazioni di batteri, dei banchi di pesci e degli stormi di pipistrelli.

Gli algoritmi di SI presentano le seguenti 3 proprietà (Dorigo e Birattari, 2007):

- Contemporaneità delle azioni: definisce la capacità degli agenti componenti lo sciame di effettuare una moltitudine di azioni in luoghi differenti e nello stesso momento. Nei sistemi artificiali questa proprietà è fondamentale, poiché consente di sviluppare sistemi più flessibili capaci di auto-organizzarsi in gruppi che considerano simultaneamente diversi aspetti di un particolare problema.
- Scalabilità: ovvero nel momento in cui si verifica un incremento della dimensione del sistema (o della quantità delle variabili), questo conserva la sua funzione senza dover riconsiderare le modalità con cui gli agenti interagiscono. Nei sistemi di SI le iterazioni coinvolgono solo gli agenti vicini tra loro, quindi il numero di iterazioni tende a non crescere proporzionalmente con la numerosità degli individui dello sciame. Grazie al principio di scalabilità un sistema può incrementare la sua performance semplicemente aumentando la sua dimensione e, soprattutto, senza la necessità di alcuna riprogrammazione.
- Tolleranza d'errore: proprietà che deriva dalla natura decentralizzata ed auto-organizzata dei sistemi di SI. Data la presenza di molti individui intercambiabili, dove nessuno di questi ha il compito di controllare il

comportamento dello sciame, ogni individuo debole può essere facilmente eliminato e sostituito da un altro pienamente funzionante. Di conseguenza, tale principio consente di oltrepassare eventuali errori senza tuttavia creare particolari problemi alla computazione.

La famiglia degli algoritmi di SI, come per gli EA, è molto vasta, e negli ultimi anni si è assistito ad un'intensificazione della ricerca in tale ambito. I 3 principali algoritmi appartenenti a questa classe sono: Ant Colony Optimization (ACO), Artificial Bee Colony algorithm (ABC) e Particle Swarm Optimization (PSO).

3.3 Particle Swarm Optimization

Come si vedrà nel successivo capitolo, il problema di allocazione proposto da *DeltaHedge* presenta una serie di vincoli a variabili miste-intere che lo rendono non lineare e non differenziabile. Data la complessità derivante dalla risoluzione di problemi di programmazione matematica NP-Hard attraverso un algoritmo esatto, a tal fine verranno utilizzate due metaeuristiche appartenenti alla famiglia delle Swarm Intelligence, ovvero la Particle Swarm Optimization (PSO) e la Bacterial Foraging Optimization (BFO). Quindi se da un lato ci si dovrà "accontentare" di una soluzione approssimata[25], dall'altro lato si potrà conseguire un notevole risparmio di costi in termini di tempo macchina e ricerca.

La PSO è una metaeuristica *population-based* proposta da Kennedy ed Eberhart nel 1995. Inizialmente fu concepita con l'obiettivo di simulare, a livello grafico, il movimento degli stormi di uccelli durante la ricerca del cibo (da qui la classificazione come *bio-inspired*). In seguito, si scoprì che si trattava anche di un ottimo metodo per la risoluzione di problemi di ottimizzazione non vincolata.

La popolazione della PSO è composta da uno sciame di particelle che si muove iterativamente all'interno dello spazio di ricerca, con lo scopo di determinare la soluzione ottima che massimizza o minimizza una determinata funzione obiettivo definita *fitness*. Analogamente a quanto detto per i sistemi di SI, la capacità risolutiva della PSO scaturisce dalle iterazioni sociali esistenti tra le particelle che formano lo

[25] In molti casi, il guadagno che si ottiene nel raggiungere la soluzione esatta non è poi così elevato rispetto all'utilizzo di un metodo approssimato.

sciame. Tale concetto è di fondamentale importanza e garantisce il collegamento tra natura ed algoritmi.

Entrando più nel dettaglio, la ricerca della soluzione ottima, attraverso la PSO, inizia dalla generazione casuale di un gruppo di soluzioni che rappresentano lo sciame di particelle dell'algoritmo. Dato che queste devono muoversi all'interno dello spazio di ricerca, ad ognuna viene associata una velocità, inizializzata a sua volta casualmente. In seguito ogni elemento dello stormo esplora una determinata zona, memorizzando e condividendo con il suo vicinato la miglior posizione raggiunta fino a quel momento. La qualità di ogni soluzione (o posizione) del problema di ottimizzazione è misurata tramite una predeterminata funzione obiettivo (cosiddetta fitness): nel momento in cui si volesse minimizzare quest'ultima, è evidente che le migliori posizioni raggiunte dalle singole particelle sono quelle a cui è associato un valore della fitness più basso. Le due forze che influenzano ad ogni iterazione (successiva alla prima) le velocità delle singole particelle sono, da un lato, il loro miglior passato e, dall'altro lato, la miglior posizione esplorata fino a quel momento dal vicinato delle stesse. Quindi è logico aspettarsi che l'intero stormo alla fine si avvicini alla posizione ottima globale (Figura 11).

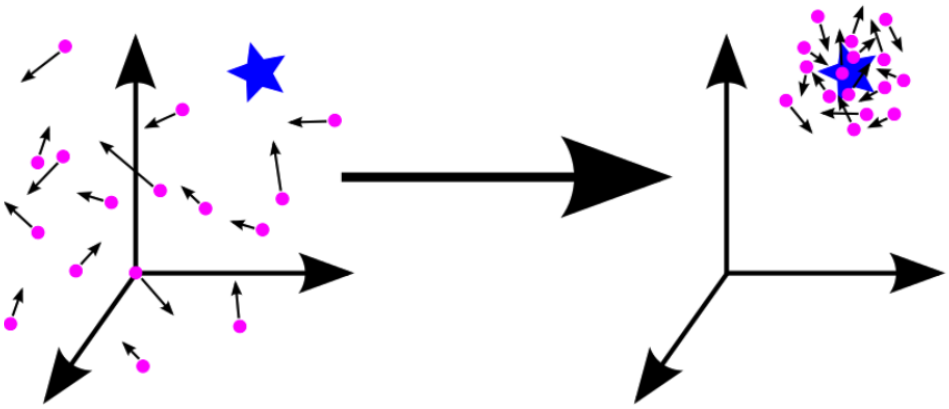


Figura 11: esempio del movimento delle particelle all'aumentare del numero di iterazioni su un piano tridimensionale.

Da un punto di vista quantitativo, ogni singola particella dello sciame costituisce un punto nello spazio *M-dimensionale*, dove *M* rappresenta il numero di variabili del problema. Inoltre, per ogni *K-iterazione* del processo, ognuna di queste è contraddistinta dai seguenti tre vettori *M-dimensionali*:

- x_j^K, ovvero la posizione della *j-esima* particella all'iterazione *k* dell'algoritmo;
- v_j^K, ovvero la velocità della *j-esima* particella all'iterazione *k* dell'algoritmo;
- p_j, ovvero la miglior posizione esplorata fino a quel momento dalla *j-esima* particella, dove il corrispondente valore della funzione di fitness è indicato con *Pbest*.

Inoltre, tutte le particelle conoscono la migliore posizione globale p_g raggiunta finora dallo sciame, il cui valore della fitness è denotato con $gbest = f(p_g)$.
Da un punto di vista di selezione di portafoglio, ogni singola particella che si trova in una determinata posizione x_j^K rappresenta un vettore contente i pesi attribuiti a ciascun titolo azionario: quindi sarà più appropriato parlare di sciami di portafogli e non di particelle. Durante le varie iterazioni dell'algoritmo tutti i portafogli vengono valutati, e nel momento in cui qualcuno di questi presentasse un valore della fitness più basso rispetto allo stesso della loro migliore posizione passata, tali nuovi portafogli verrebbero salvati nei rispettivi vettori p_j. L'obiettivo di questo processo è proprio quello di determinare nuove posizioni che siano in grado di perfezionare il più possibile la funzione obiettivo.
Il notevole successo che ha avuto la PSO negli ultimi anni è giustificato dalla sua facilità di implementazione e computazione. Essendo una metaeuristica di tipo *derivative-free*, si possono utilizzare differenti funzioni di fitness (anche non differenziabili) senza imbattersi nel problema del calcolo delle derivate prime o seconde in fase di ottimizzazione. Inoltre è estremamente flessibile, infatti può essere utilizzata per risolvere problemi appartenenti a una moltitudine di ambiti diversi semplicemente variando la tipologia della funzione obiettivo.
Nonostante i vantaggi appena definiti, si rende necessario ribadire che la PSO è una metaeuristica e, di conseguenza, non è garantito il raggiungimento della soluzione ottima globale del problema che si sta analizzando: si potrebbe ottenere un ottimo locale e confondere quest'ultimo come quello globale. Inoltre, come evidenziato da Blackwell *at al.* (2007), dato che le velocità delle particelle sono influenzate solamente dalla loro migliore posizione passata e dalla miglior posizione del loro vicinato, è evidente che una parte dell'informazione rimane non utilizzata durante l'intero processo: tutte le posizioni passate visitate dalle particelle, ma non valutate come migliori, sono informazioni che rimangono inutilizzate.

Dopo questa breve introduzione, nei successivi due paragrafi si entrerà più nel dettaglio della PSO presentando prima l'algoritmo originale, proposto da Kennedy ed Eberhart, e poi le principali modifiche introdotte nel corso degli anni.

3.3.1 L'algoritmo originale

L'algoritmo originale della PSO si esplicita nei seguenti semplici passaggi:

1. Generazione iniziale delle particelle che compongono lo stormo, attribuendo a ciascuna di queste posizione e velocità entrambe casuali.
2. Avvio ciclo *for*:
   2.1. Determinare il valore della funzione di fitness $f(x_j^k)$ per ogni posizione attuale x_j^k in cui si trova ogni singola particella;
   2.2. Se $f(x_j^k)$ è migliore del *Pbest* si sostituisce la migliore posizione passata p_j con quella corrente x_j^k;
   2.3. Selezionare le particelle del vicinato che si trovano nella miglior posizione ed attribuirle l'indice p_g (punto di ottimo globale);
   2.4. Determinare la nuova posizione x_j^{k+1} e velocità V_j^{k+1} di ogni particella utilizzando le seguenti equazioni:

$$\begin{cases} V_j^{k+1} = V_j^k + U(0,\phi_1) \otimes \left(p_j - x_j^k\right) + U(0,\phi_2) \otimes (p_g - x_j^k) \\ x_j^{k+1} = x_j^k + V_j^{k+1} \end{cases},$$

   dove:
   - $U(0,\phi_i)$ è un vettore di numeri casuali uniformemente distribuiti nell'intervallo $[0,\phi_i]$;
   - ϕ_1 e ϕ_2 sono i coefficienti di accelerazione;
   - $\otimes$ è il prodotto tensoriale.

   2.5. Se si raggiunge un predeterminato criterio di stop, che può essere un soddisfacente valore della funzione di fitness o un limite massimo di iterazioni, si passa al punto 3, altrimenti si torna al punto 2 dopo aver aggiornato il numero delle iterazioni.
3. Fine del ciclo.

Il punto fondamentale dell'intero processo è senza dubbio il calcolo della velocità

all'iterazione *k+1*. Questa, come si può notare, è costituita da tre addendi:

- Il primo rappresenta la velocità corrente V_j^k della *j-esima* particella;
- Il secondo, ovvero la differenza tra la migliore posizione passata e quella attuale, esprime la parte cognitiva del comportamento della particella;
- Il terzo, ossia la differenza tra la migliore posizione globale e quella attuale, esprime la parte sociale del comportamento della particella.

Quindi si può affermare che la velocità futura di una particella è pari alla triplice somma tra la sua velocità passata, una correzione verso la sua migliore posizione visitata fino a quel momento e una correzione verso la miglior posizione del suo vicinato. Il secondo e il terzo addendo vengono entrambi moltiplicati per un vettore di numeri casuali poiché si cerca di disturbare il comportamento della particella, in modo da spingerla ad approfondire la ricerca nelle zone limitrofe.
I due coefficienti di accelerazione ϕ_1 e ϕ_2 influenzano in maniera determinante l'intensità delle forze che attraggono ciascuna particella verso p_j o p_g. Un loro settaggio errato può comportare un aumento incontrollato delle velocità delle particelle (instabilità del processo). Inizialmente, la prassi comune era quella di utilizzare un valore di ϕ_1 e ϕ_2 pari a 2, che non garantiva la stabilità della PSO. Di conseguenza, per superare tale limite, si propose di confinare la velocità V_j^k all'interno di un determinato range di valori $[-V_{max}, +V_{max}]$. Sebbene tale soluzione comporti un contenimento dell'instabilità del processo, dall'altro lato non è ben chiaro quale debba essere il valore appropriato di V_{max}.
Un secondo parametro che deve essere fissato per il corretto funzionamento della PSO è il numero di particelle componenti lo stormo; un numero troppo elevato richiederebbe notevoli sforzi computazionali. A livello empirico la prassi è quella di utilizzare uno stormo composto dai 20 ai 50 componenti (Blackwell *at al.*,2007).

3.3.2 Varianti della PSO

Peso d'inerzia

Una prima variante dell'algoritmo originale della PSO è quella proposta da Shi ed Eberhart nel 1998, ovvero l'*Inertia Weight*. Questa modifica cerca di migliorare la potenzialità risolutiva della PSO, con il secondo fine di limitare o (in altri casi)

eliminare l'utilizzo del parametro V_{max}. In particolare, l'equazione che definisce il valore della velocità futura di ogni singola particella, si modifica nel seguente modo:

$$\begin{cases} V_j^{k+1} = \omega V_j^k + U(0,\phi_1) \otimes \left(p_j - x_j^k\right) + U(0,\phi_2) \otimes (p_g - x_j^k) \\ x_j^{k+1} = x_j^k + V_j^{k+1} \end{cases},$$

dove ω è il peso d'inerzia.

La funzione di tale parametro è quella di bilanciare i due fondamentali comportamenti della PSO, ovvero lo sfruttamento e l'esplorazione: il primo è determinato dagli ultimi due addendi, mentre il secondo è specificato dalla componente ωV_j^k. È evidente che all'aumentare del valore attribuito a ω, più elevata sarà l'influenza della velocità corrente su quella futura. All'opposto, un valore del peso d'inerzia basso conduce ad un aumento della dipendenza sia dalla migliore posizione personale della particella, sia da quella globale di tutto il vicinato. Un metodo che permette di sfruttare il peso d'inerzia in modo razionale è quello di attribuirgli un valore decrescente all'aumentare del numero di iterazioni. Difatti mentre nelle fasi iniziali è preferibile rafforzare l'esplorazione (con un ω elevato), una volta che viene raggiunta una zona "attraente" nello spazio delle soluzioni è opportuno approfondire la ricerca mediante l'opera di sfruttamento (con un ω più basso). Una possibile formulazione per determinare il valore del peso d'inerzia all'aumentare delle iterazioni dell'algoritmo è quella proposta da Corazza *at al*. (2013):

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

dove:

- k rappresenta il numero dell'iterazione attuale;
- K è il numero massimo di iterazioni consentite.

Per quanto riguarda i valori numerici attribuiti a ω_{max} e ω_{min}, la prassi è quella di settarli rispettivamente pari a 0.9 e 0.4. Gli stessi autori evidenziano come ai coefficienti di accelerazione ϕ_1 e ϕ_2 debbano essere assegnati dei valori coerenti con quelli attribuiti ai pesi d'inerzia.

In letteratura si possono trovare altri metodi per "correggere" ω^k. Ad esempio, Shi ed Eberhart (2000) propongono un adattamento di ω^k utilizzando un sistema *fuzzy*.

Ci sono inoltre altri studi, come Zheng *et al*. (2003), nei quali si ottengono buoni risultati utilizzando un peso d'inerzia crescente.

Constriction coefficients

Una seconda variante della PSO ampiamente utilizzata è la "limitazione dei coefficienti", introdotta da Clerc e Kennedy nel 2002. Questa, oltre a prevenire l'esplosione della velocità, permette di garantire la convergenza delle particelle ed eliminare il parametro arbitrario V_{max}. Il calcolo della velocità futura si modifica nel seguente modo:

$$\begin{cases} V_j^{k+1} = \chi\left(V_j^k + U(0,\phi_1) \otimes \left(p_j - x_j^k\right) + U(0,\phi_2) \otimes \left(p_g - x_j^k\right)\right) \\ x_j^{k+1} = x_j^k + V_j^{k+1} \end{cases},$$

dove:

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

Per quanto riguarda il valore di ϕ, la prassi è quella di porlo pari a 4.1 con $\phi_1 = \phi_2$; da questo ne consegue un valore di χ pari a 0.7298.

Diversamente da quanto visto nel caso del peso d'inerzia, il coefficiente di costrizione χ è applicato su tutti i componenti dell'equazione della velocità futura. Quindi, mentre V_j^k è moltiplicato per 0.7298, gli altri due addendi, che determinano lo sfruttamento, sono moltiplicati per un valore casuale all'interno dell'intervallo [0,1.496118][26].

Naturalmente esiste una connessione tra questo metodo e quello definito precedentemente: questi sono algebricamente equivalenti quando $\omega = \chi$ e $\phi_i = \phi_i \chi$.

Fully Informed Particle Swarm (FIPS)

Il metodo FIPS, proposto da Kennedy e Mendes nel 2002, ha l'obiettivo di sfruttare l'informazione inutilizzata che si trova nel vicinato delle singole particelle. Quindi, mentre nell'algoritmo originale ogni particella è influenzata solamente dalla

[26] Dato che $\phi_1 = \phi_2 = 2.05$ allora $0.7298 * 2.05 \cong 1.49618$.

migliore posizione personale e dalla migliore posizione di tutto il vicinato, in questo caso le particelle dipendono dalle posizioni di tutti i loro vicini. Il FIPS si sviluppa nel seguente modo:

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

dove:

- K_j definisce il numero di vicini della *j-esima* particella;
- nbr_n^j rappresenta l'*n-esimo* vicino della *j-esima* particella.

Come si può notare, i risultati ottenuti mediante questo modello coincidono con gli stessi ricavati attraverso l'algoritmo originale solamente quando K_j è pari a 2. Kennedy e Mendes mostrano che scegliendo dei parametri adeguati per il FIPS, è possibile determinare delle soluzioni migliori in un numero di iterazioni inferiore rispetto alla PSO originale.

3.3.3 Topologia della popolazione

Fino a questo punto si è sempre parlato di "vicinato" delle particelle senza mai specificare come questo sia strutturato. Tuttavia, la scelta della topologia dell'algoritmo è di fondamentale importanza in quanto definisce le modalità con cui le particelle si scambiano le informazioni tra loro.

A livello macro, tutte le topologie possono essere di due tipi: statiche o dinamiche. Le prime sono caratterizzate dal fatto che la struttura del vicinato delle particelle rimane sempre la stessa al variare delle iterazioni. In particolare, le principali topologie appartenenti a quest'ultimo gruppo sono:

- Local best (lbest) proposta da Eberhart e Kennedy (1995). Questa presenta una semplice struttura ad anello (Figura 12), dove ogni particella comunica con i suoi due vicini adiacenti. Il principale vantaggio di tale topologia è quello di generare sub-popolazioni multiple, che permettono di effettuare ricerche parallele nello spazio delle soluzioni. Da questo ne deriva una minore vulnerabilità all'attrazione verso un ottimo locale in confronto ad altre topologie (come la gbest).

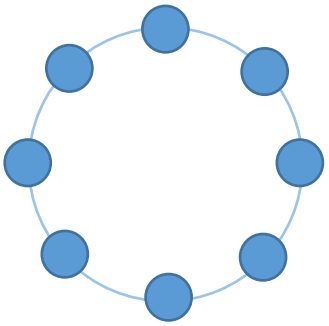


Figura 12: topologia lbest.

- Global best (gbest), dove ogni particella è connessa con tutte le altre componenti lo stormo (Figura 13). In altri termini è un sistema *fully connected* nel quale ciascuna particella scambia informazioni con tutto lo sciame; quindi non si generano sub-popolazioni come nel caso precedente.

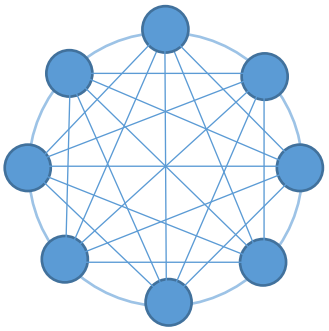


Figura 13: topologia gbest.

- Topologia Von Neumann, che può essere considerata una via di mezzo tra la gbest e la lbest. In questo caso ciascun componente dello sciame è connesso sia con le particelle che si trovano alla sua destra e alla sua sinistra, sia con quelle che si posizionano sopra e sotto dello stesso (Figura 14). Kennedy e Mendes (2002) hanno mostrato la superiorità di tale topologia rispetto alle precedenti.

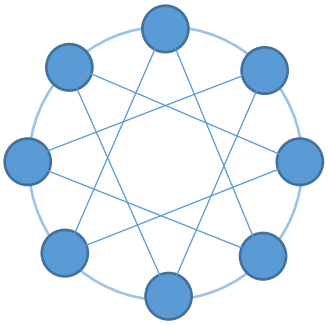


Figura 14: topologia Von Neumann.

Per quanto riguarda la seconda macro-classe, ovvero le topologie dinamiche, queste sono contraddistinte da una struttura del vicinato che si modifica all'aumentare del numero di iterazioni. Le principali topologie di questo tipo che si possono trovare in letteratura sono le seguenti:

- Suganthan (1999), partendo dal presupposto che la lbest è migliore nella fase di esplorazione mentre la gbest converge più velocemente, propose di iniziare la ricerca con la lbest per poi incrementare lentamente l'ampiezza del vicinato finché non viene raggiunta una topologia di tipo gbest alla fine dell'algoritmo.
- Peram *et al.* (2003) proposero di utilizzare una distanza Euclidea pesata, con il fine di identificare la particella con cui dovrebbe interagire ogni componente dello sciame. Più in dettaglio, tutte le particelle selezionano il proprio vicino sulla base di un "fitness distance ratio" (FDR):

$$FDR = \frac{Fitness(X_j) - Fitness(P_i)}{|X_j - P_i|},$$

  dove P_i è l'*i-esimo* vicino della particella *j-esima*. Questo rapporto garantisce una diminuzione della probabilità che una particella interagisca con un'altra che si trova in una zona molto più lontana rispetto alla prima, all'interno dello spazio di ricerca.
- Liang e Suganthan (2005) proposero di generare casualmente diverse sub-popolazioni di dimensione *n*, con delle connessioni tra particelle che variano, a loro volta casualmente, ogni *k* iterazioni.

- Jason e Middenford (2005) proposero una topologia di tipo "gerarchica", dove ogni particella appartenente allo sciame è soggetta ad una duplice influenza da parte sia della sua migliore posizione passata, sia della particella che si posiziona ad un livello immediatamente superiore per quanto riguarda la qualità del valore della fitness.
- Clerc (2006) introdusse la topologia "Tribes", nella quale i parametri che la caratterizzano, come ad esempio la dimensione della popolazione, evolvono nel corso delle iterazioni a seconda delle performance ottenute. In particolare, dopo aver generato diverse sub-popolazioni, si selezionano le particelle peggiori che si trovano all'interno delle migliori sub-popolazioni, con il fine di spostarle all'interno delle altre sottopopolazioni.

3.4 Bacterial Foraging Optimization

La Bacterial Foraging Optimization (BFO), proposta da Passino nel 2002, è una delle più recenti metaeuristiche *bio-inspired* che si possono trovare in letteratura. Sin dalla sua introduzione, questa ha attirato l'attenzione di molti ricercatori provenienti da diversi ambiti, grazie alla sua elegante struttura e, soprattutto, alla sua efficienza nel risolvere problemi di ottimizzazione reali.

La BFO, come si può intuire dal nome, è ispirata dai movimenti di un batterio piuttosto comune presente nell'intestino dell'uomo, ovvero l'*Escherichia Coli* (E. coli). Quest'ultimo è costituito da un diametro di 1 μm, una lunghezza di circa 2 μm e, sotto appropriate condizioni, può riprodursi (dividersi) in 20 minuti. La locomozione è garantita da una serie di specifiche appendici cellulari, dette flagelli[27], che hanno lo scopo di aiutare il batterio ad eseguire i due movimenti di base che lo caratterizzano: l'avanzamento (ai 10-20 μm/sec), quando l'E. coli nuota in una determinata direzione in modo lineare; il capovolgimento, quando l'E. coli si ferma e cambia direzione. In particolare, il primo movimento è garantito dalla rotazione in senso antiorario dei flagelli. Quando invece quest'ultimi si muovono in direzione

[27] I flagelli sono degli organi di locomozione tipici dei batteri a forma cilindrica (bacilli). Questi hanno una struttura filamentosa e sono costituiti da subunità proteiche elicoidali contenenti flagellina. Ogni flagello è costituito da 3 parti: il filamento, ovvero la porzione sporgente; un gancio, mediante il quale si attacca alla membrana plastica; un corpo basale, che funge da ancoraggio alla membrana.

oraria, i batteri cessano di nuotare e si capovolgono.
Qualora nell'ambiente circostante fosse presente un attrattivo, l'E. coli eseguirebbe meno "capriole" con il fine di restare più tempo possibile nell'area attraente, mentre, in presenza di repellenti, il numero dei capovolgimenti aumenterebbe. Quindi un batterio potrebbe nuotare nella stessa direzione per molto tempo se la concentrazione nutriente attorno a lui stesse incrementando. Tuttavia, nel momento in cui si accorgesse che la zona ricca di nutrienti sta diminuendo, cambierebbe velocemente direzione mediante un capovolgimento. Si può affermare che l'E. coli si sposta secondo gradiente, ovvero intensifica le "capriole" sia quando diminuisce il gradiente di concentrazione dell'attrattivo, sia nel momento in cui aumenta il gradiente del repellente. Tali movimenti caratterizzano l'attività di chemiotassi del batterio, ossia il propulsore chimico della struttura di risposta sensoriale che condiziona la funzione dei flagelli.
Da un punto di vista quantitativo, la chemiotassi può essere espressa con la seguente formula:

$$\theta^i(j+1,k,l) = \theta^i(j,k,l) + C(i) * \phi(j),$$

dove:

- $\theta^i(j,k,l)$ è la posizione del batterio i alla *j-esima* iterazione chemiotattica, *k-esima* iterazione di riproduzione e *l-esima* iterazione di eliminazione-dispersione;
- $C(i)$ è la dimensione della fase di chemiotassi dell'*i-esimo* batterio;
- $\phi(j)$ è un numero casuale che ha lo scopo di definire la direzione dei movimenti in caso di capovolgimento del batterio alla *j-esima* iterazione chemiotattica.

Dopo aver eseguito reiteratamente i movimenti di chemiotassi per un determinato periodo di tempo, lo stormo di batteri proverà a riprodursi con l'obiettivo di migliorare le sue probabilità di sopravvivenza. A tal scopo ciascuno dei batteri più sani, ovvero quelli che possiedono un miglior valore della fitness, si divide asessualmente (scissione binaria) a metà, generando due nuovi batteri che si posizionano nella stessa zona del loro "genitore". Inoltre, nello stesso momento in cui avviene tale riproduzione, i batteri più deboli muoiono affinché la dimensione dello stormo rimanga costante. È evidente che mediante questa operazione i batteri

più forti sostituiscono la maggior parte dei batteri più deboli. Questo concetto è molto simile al principio del *survival of the fittest* visto negli algoritmi evolutivi.
A causa di improvvisi o graduali cambiamenti dell'ambiente dove si trova uno stormo di batteri, si potrebbe verificare l'evento di eliminazione e dispersione: tutti i batteri che si trovano in una determinata regione potrebbero morire oppure essere dispersi in una nuova zona dell'ambiente. In caso di eliminazione, nuovi batteri sono generati e posizionati su un'area casuale dello spazio di ricerca. Questo processo è molto simile all'operazione di mutazione che si verifica negli algoritmi genetici.
Sebbene la probabilità dell'evento di eliminazione e dispersione sia molto bassa per ogni singolo batterio, tale procedimento, considerando un ampio arco temporale, può incrementare la diversità delle soluzioni, garantire la capacità di ricerca globale ed evitare di rimanere intrappolati in ottimi locali (Kao e Cheng, 2013).
Le cellule, quando stimolate da un elevato livello di *succinato*[28], rilasciano un attraente (*aspartato*), che le aiuta ad aggregarsi in gruppi e, di conseguenza, a muoversi come dei modelli concentrici di sciami con un'elevata densità di batteri. Nel caso degli E. coli, tali segnali, che generano una sorta di comportamento "sociale", possono essere modellizzati mediante la seguente funzione:

$$J_{cc}\big(\theta, P(j,k,l)\big) =$$
$$= \sum_{i=1}^{S} J_{cc}\left(\theta, \theta^i(j,k,l)\right) = \sum_{i=1}^{S}[-d_{attractant}\exp(-w_{attractant}\sum_{m=1}^{p}(\theta_m - \theta_m^i)^2)] +$$
$$+ \sum_{i=1}^{S}[h_{repellant} exp\,(-w_{repellant}\sum_{m=1}^{p}(\theta_m - \theta_m^i)^2)],$$

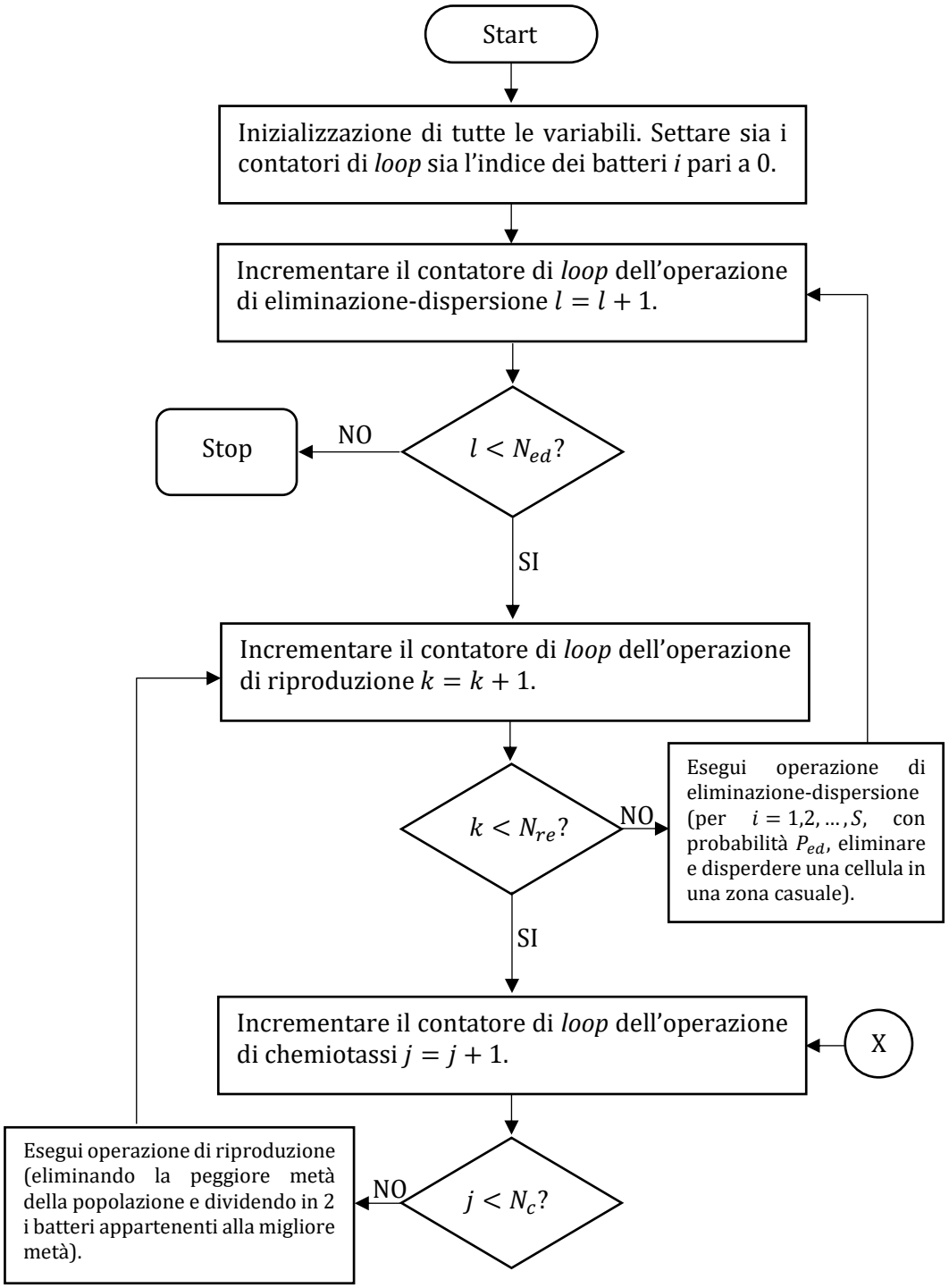
dove:

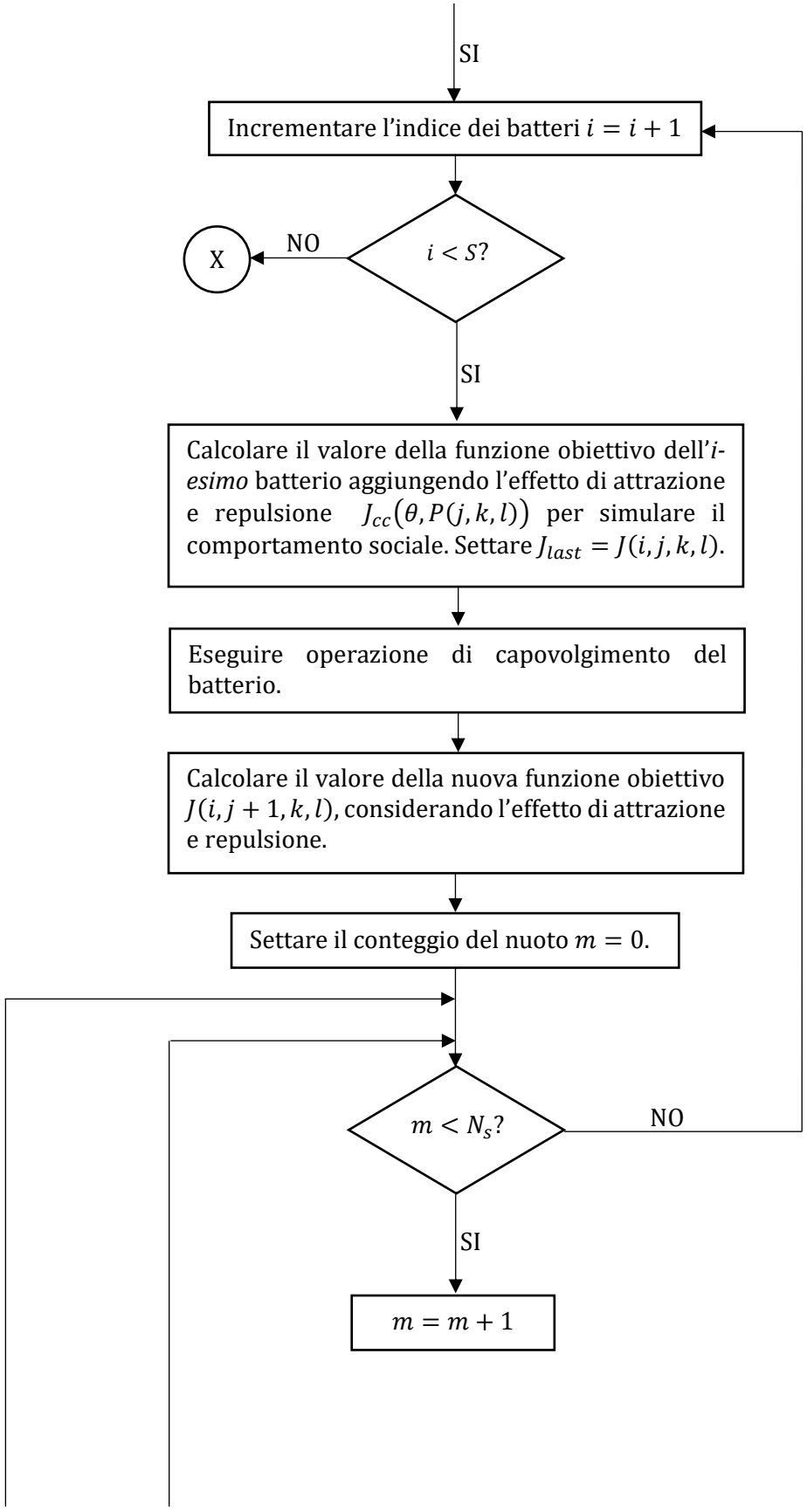
- $J_{cc}\big(\theta, P(j,k,l)\big)$ è il profilo attrattivo-repellente tra le cellule che simula il comportamento sociale[29];
- S è il numero totale di batteri;
- p è il numero di variabili da ottimizzare e che sono presenti in ogni batterio;
- $\theta = (\theta_1, \theta_2, \dots, \theta_p)$ è un punto nel dominio *p-dimensionale;*
- $d_{attractant}$, $w_{attractant}$ sono due coefficienti che rappresentano rispettivamente l'intensità e l'ampiezza dei segnali attrattivi rilasciati dai batteri;

[28] Il succinato è un acido dicarbossilico sintetizzato nel ciclo di Krebs e si attiva per la reazione del succinil-CoA che agisce con l'enzima succinil-CoA sintetasi.
[29] Questo valore deve essere aggiunto all'effettiva funzione obiettivo (da minimizzare).

- $h_{repellant}$, $w_{repellant}$ sono due coefficienti che rappresentano rispettivamente l'intensità e l'ampiezza dell'effetto repellente.

Per meglio illustrare il funzionamento di tale algoritmo si veda il seguente diagramma di flusso:




SI
Incrementare l'indice dei batteri $i = i + 1$
NO
X
$i < S$?
SI
Calcolare il valore della funzione obiettivo dell'*i-esimo* batterio aggiungendo l'effetto di attrazione e repulsione $J_{cc}(\theta, P(j,k,l))$ per simulare il comportamento sociale. Settare $J_{last} = J(i,j,k,l)$.
Eseguire operazione di capovolgimento del batterio.
Calcolare il valore della nuova funzione obiettivo $J(i, j+1, k, l)$, considerando l'effetto di attrazione e repulsione.
Settare il conteggio del nuoto $m = 0$.
$m < N_s$?
NO
SI
$m = m + 1$

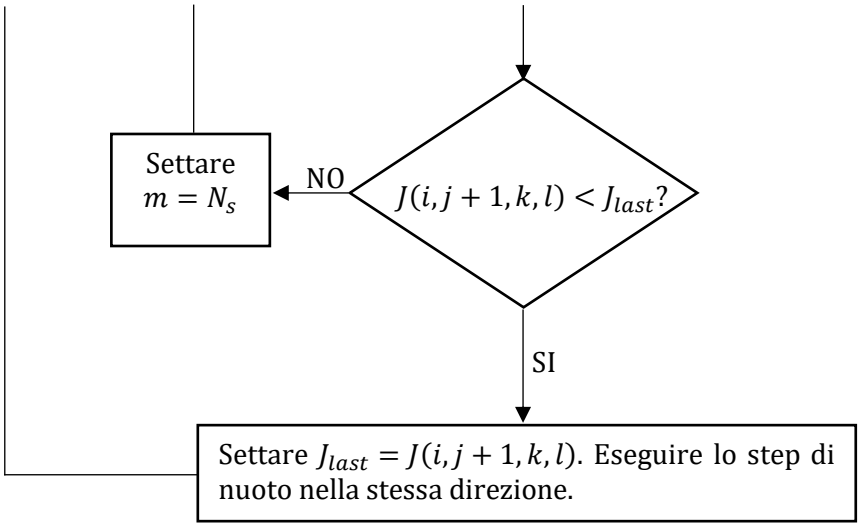


Figura 15: diagramma di flusso della BFO.

dove N_{ed} è il numero totale degli eventi di eliminazione, N_{re} è il numero totale delle fasi di riproduzione e N_c è il numero totale di movimenti chemiotattici.

Lo scopo di questa breve introduzione sulla BFO era quello di definire il collegamento esistente tra le principali fasi biologiche degli E. coli e i passi dell'algoritmo originale ispirato ai movimenti di tali batteri. Nel successivo paragrafo, seguendo quanto proposto da Kao e Cheng (2013), si presenterà un approccio per la risoluzione dei problemi di ottimizzazione di portafoglio mediante l'algoritmo BFO.

3.4.1 BFO per la risoluzione di problemi di ottimizzazione di portafoglio

Da un punto di vista di ottimizzazione di portafoglio, ogni singolo batterio artificiale rappresenta un portafoglio, ovvero un vettore contenente i pesi attribuiti a ciascun titolo. La determinazione della soluzione ottima mediante tale algoritmo può essere definita come la combinazione di 3 operazioni:

1. *Chemiotassi*: il principale compito di tale step è quello di cercare migliori soluzioni attraverso i movimenti (avanzamento e capovolgimento) dei batteri. Durante uno step di chemiotassi ogni batterio esegue prima un capovolgimento e, successivamente, un certo numero di avanzamenti. Eseguendo una moltitudine di tali step, i batteri artificiali producono nuove soluzioni attraverso l'aumento e la diminuzione dei pesi associati ad ogni

titolo. In altri termini, considerando una popolazione di portafogli B, l'algoritmo aggiusta le quote (xx_{bi}) attribuite ad ogni titolo componente il portafoglio mediante la duplice fase di capovolgimento ed avanzamento. Dopo aver generato una nuova soluzione (nxx_{bi}), l'algoritmo la valuta mediante una determinata funzione obiettivo e la sostituisce alla precedente nel caso fosse migliore. L'equazione di chemiotassi può essere definita nel seguente modo:

$$nxx_{bi} = xx_{bi}^{t} + C(i) * \Delta D_b(i), \forall i$$

$$xx_{bi}^{t+1} = \begin{cases} nxx_{bi}\ se\ f(ns_b) < f(s_b) \\ xx_{bi}^{t}\ in\ caso\ contrario \end{cases} \forall i$$

dove:

- $f(s_b)$ e $f(ns_b)$ sono rispettivamente il valore della funzione obiettivo della soluzione corrente e della nuova soluzione;
- $C(i)$ è un parametro che rappresenta l'ampiezza dello step di chemiotassi dell'*i-esimo* batterio;
- $\Delta D_b(i)$ è un numero casuale ($\in [-1,1]$) che definisce l'ammontare del cambiamento di direzione dell'*i-esimo* batterio dovuto alla fase di capovolgimento;
- t è il numero degli step di chemiotassi.

Il parametro $C(i)$ differisce sulla base della tipologia dello step di chemiotassi: $C_t(i)$ indica la dimensione dello step di capovolgimento e $C_s(i)$ si riferisce alla dimensione di uno step di avanzamento. Ogni fase di chemiotassi include un capovolgimento e N_s avanzamenti. Per rendere l'algoritmo più efficiente, tutti i batteri utilizzano un meccanismo di autoverifica per controllare i due movimenti che lo caratterizzano. Se un capovolgimento (o un avanzamento) non produce un risultato migliore, il batterio ne esegue immediatamente un altro per cambiare la sua direzione. Quindi, per ogni asset selezionato, un batterio prima realizza un movimento di capovolgimento per determinare il nuovo valore di $\Delta D_b(i)$, poi calcola il nuovo valore di xx_{bi} (nxx_{bi}) mediante l'equazione precedente. Quando il valore di xx_{bi} diventa negativo ($nxx_{bi} < 0$) l'algoritmo compie una procedura di riparazione, nella quale l'asset i è sostituito con un asset non

selezionato. Nel momento in cui il capovolgimento fornisce un risultato migliore, prima si rimpiazza la soluzione corrente s_b con la nuova soluzione ns_b, poi si eseguono N_s avanzamenti. Tuttavia, se il risultato derivante dal capovolgimento non è accettabile, il batterio deve fare nuovi capovolgimenti finché non trova una soluzione migliore o fino a quando non viene raggiunto il numero massimo di step di chemiotassi (N_s) durante uno step di riproduzione. Dopo un'operazione di capovolgimento eseguita con successo, il batterio effettua i movimenti di avanzamento, che possono essere di un numero di iterazioni inferiore a N_s nel caso in cui non fosse più possibile migliorare ulteriormente la soluzione. In ogni step di avanzamento, si determina un nuovo valore di xx_{bi} utilizzando lo stesso $\Delta D_b(i)$ ricavato dal precedente capovolgimento. Come nel caso precedente, se tale step offre una soluzione migliore, la BFO esegue un'operazione di sostituzione.

2. *Riproduzione*: dopo N_c fasi di chemiotassi si esegue uno step di riproduzione, che ha il compito principale di generare eventuali migliori soluzioni per aumentare efficientemente la qualità di tutta la popolazione. Il primo passo di tale step è quello di classificare i membri della popolazione in ordine ascendente sulla base dei valori delle loro funzioni obiettivo. La prima metà della popolazione è selezionata per la riproduzione mentre la seconda metà è scartata. Da ogni membro selezionato si ottengono due nuovi batteri, che vengono piazzati nella stessa posizione del primo: in questo modo la dimensione della popolazione non cambia. Dopo uno step di riproduzione si eseguono altre N_c fasi di chemiotassi. La BFO ripete iterativamente quest'ultimo ciclo finché il numero degli step di riproduzione non raggiunge N_{re}.
3. *Eliminazione-dispersione*: dopo N_{re} fasi di riproduzione si esegue uno step di eliminazione-dispersione, che ha il compito principale di incrementare la potenziale diversità della popolazione. Ogni asset componente una determinata soluzione ha una probabilità pari a P_{ed} di subire tale step, ovvero di essere eliminato e rimpiazzato da un asset non selezionato precedentemente. Dopo questa operazione, la BFO esegue N_{re} step di riproduzione, dove in ognuno dei quali vengono realizzati N_c step di

chemiotassi in modo ripetitivo. Questo ciclo viene ripetuto più volte finché il numero delle fasi di eliminazione-dispersione non raggiunge il valore N_{ed}.

In conclusione, si presenta uno pseudo-codice della BFO per la risoluzione di un problema di ottimizzazione di portafoglio (Kao e Cheng, 2013).

Pseudo-codice BFO

Inizializzazione in modo casuale di tutti i batteri
Calcola $f_b, b = 1, \dots, B$ 'Calcola il valore della funzione obiettivo di ogni batterio'
per $ned = 1$ **fino a** N_{ed} **fai** 'N_{ed}: numero massimo di fasi di eliminazione-dispersione'
  per $nre = 1$ **fino a** N_{re} **fai** 'N_{re}: numero massimo di fasi di riproduzione'
    per $nc = 1$ **fino a** N_c **fai** 'N_c: numero massimo di fasi di chemiotassi'
    '**step di chemiotassi**'
      per $b = 1$ **fino a** B **fai**
        per $i = 1$ **fino a** K **fai** 'K: numero dei titoli selezionati'
          esegui movimento di capovolgimento
        fine ciclo per
        Calcola $f_b^{dopoCapovolgimento}$
        se $f_b^{dopoCapovolgimento}$ è migliore **allora**
          muovi il batterio b sulla nuova posizione
          $f_b = f_b^{dopoCapovolgimento}$
          $H = H \cup s_b$ 'H: insieme di tutte le soluzioni migliorate'
        fine se
        se $f_b^{dopoCapovolgimento}$ è migliore **allora** 'inizia lo step di avanzamento'
          $n_s = 0$ 'n_s: numero di iterazione degli step di avanzamento'
          fai
            per $i = 1$ **fino a** K **fai**
              esegui step di avanzamento
            fine ciclo per
            Calcola $f_b^{dopoAvanzamento}$
            se $f_b^{dopoAvanzamento}$ è migliore **allora**
              muovi il batterio b sulla nuova posizione
              $f_b = f_b^{dopoAvanzamento}$
              $H = H \cup s_b$
            fine se
            $n_s = n_s + 1$
          mentre ($f_b^{dopoAvanzamento}$ non è migliore oppure $n_s > N_s$) 'N_s: numero massimo di fasi di avanzamento'
        fine se 'fermare gli step di avanzamento'
      fine ciclo per 'fine di uno step di chemiotassi'
      aggiorna $f_{migliore-finora}$ e $s_{migliore-finora}$
    fine ciclo per 'fine di N_c step di chemiotassi'
  '**step di riproduzione**'

ordina i batteri in base al valore di f_b 'ordina i batteri in ordine ascendente'
elimina i batteri nella seconda metà della popolazione
riproduci i batteri nella prima metà della popolazione
posiziona i batteri nella stessa zona dei loro genitori
fine ciclo per 'fine di N_{re} step di riproduzione'
'**step di eliminazione-dispersione**'
per $b = 1$ **fino a** B **fai**
per $i = 1$ **fino a** K **fai**
se $(r = random(0,1) < P_{ed})$ **allora** 'P_{ed}: probabilità di eliminazione-dispersione'
esegui step di eliminazione-dispersione
fine ciclo per
Calcola $f_b^{dopoE-D}$
se $f_b^{dopoE-D}$ è migliore **allora**
muovi il batterio b sulla nuova posizione
$f_b = f_b^{dopoE-D}$
$H = H \cup s_b$
fine se
fine ciclo per 'fine di uno step di eliminazione-dispersione'
aggiorna $f_{migliore-finora}$ e $s_{migliore-finora}$
fine ciclo per 'fine di N_{ed} step di eliminazione-dispersione'
registra il miglior batterio 'fine del ciclo della BFO'
Fine

CAPITOLO 4

IL CASO DELTAHEDGE

In questo capitolo verrà formulato il problema di ottimizzazione vincolata avanzato da DeltaHedge. Nella prima parte si presenterà nel dettaglio la società, per poi definire i principali strumenti finanziari mediante i quali opera, ovvero i futures. In seguito, si tradurranno in termini quantitativi le richieste avanzate da DeltaHedge per la determinazione dell'allocazione ottima all'interno del programma Styx, ottenendo, in questo modo, un problema di minimizzazione vincolata. Infine, dato che le metaeuristiche presentate nel capitolo precedente sono applicabili solamente a problemi di ottimizzazione libera, si riformulerà il problema di selezione iniziale in forma non vincolata attraverso il metodo delle penalità esatte.

4.1 La società

Come già accennato nei capitoli precedenti, l'obiettivo di tale elaborato consiste nel risolvere un particolare problema di ottimizzazione vincolata che mi è stato proposto durante lo stage svolto in DeltaHedge. Come si vedrà nei successivi paragrafi, tale problema di selezione si discosta dai classici problemi di allocazione di un determinato capitale in un portafoglio composto da azioni (e, in altri casi, da un titolo *risk free*), tuttavia la logica sottostante non cambia.

DeltaHedge LP., fondata a Padova nel 2011 da Umberto Loschi e Robert Baron, è una CTA (Commodity Trade Association)[30] di modello anglosassone qualificata

[30] Una Commodity Trade Association (CTA) è una società che offre consulenza personalizzata riguardo all'acquisto ed alla vendita di contratti futures, opzioni su futures e/o swaps. Le tipologie di sottostanti, sui quali si basano i contratti derivati utilizzati, possono variare dai beni fisici (come prodotti agricoli, metalli ed energia) ai strumenti finanziari (come valute, indici e bonds). I tre principali stili d'investimento eseguiti dalle CTAs sono: tecnico, quando il fine è quello di prevedere determinati trend dei prezzi mediante analisi tecniche dei grafici; fondamentale,

nell'advisory operativa indirizzata ad una clientela di elevato profilo. Più in dettaglio, tale società è specializzata nella gestione sistematica di contratti futures e nella consulenza in ambito risk management (in particolare per quanto riguarda la copertura del rischio di cambio, di tasso ed operativo): la sua operatività si focalizza principalmente nei mercati valutari e nei mercati delle commodities. Attualmente DeltaHedge offre tre differenti programmi di trading:

- *DH Nix Alpha Programs*: questo, sviluppato nel novembre del 2015, mira a generare interessanti rendimenti assoluti investendo in un solo mercato, ovvero quello dei E-mini 5$ Dow Jones Industrial Average Future scambiati nel CME Group (Chicago Mercantile Exchange & Chicago Board of Trade). Tale programma si fonda su un sistema di trading automatico che sfrutta quattro diverse strategie focalizzate, a livello quantitativo, sui movimenti intraday (con segnali di 5 e 30 minuti) dei prezzi e delle volatilità dei rendimenti. Dopo aver analizzato l'indice sottostante, il programma genera dei segnali per ogni strategia in modo indipendente. La bassa correlazione dei rendimenti generati dalle quattro strategie garantisce la riduzione del rischio di portafoglio. Il programma è inoltre integrato con un innovativo modello matematico "anti-noise" che funziona in parallelo con tali strategie di trading nel Dow Jones.
- *VixVox Program*: sviluppato nell'agosto del 2015 da Edoardo Ciravegna. L'obiettivo primario di tale programma è quello di incrementare il capitale degli investitori mediante la speculazione sugli indici di volatilità dei principali mercati azionari. Più in dettaglio, la strategia effettua posizioni giornaliere lunghe e corte sui contratti CBOE Vix Futures ed Eurex VStoxx Futures. Il posizionamento viene elaborato automaticamente sulla base degli

---

quando i *traders* tentano di prevedere i prezzi analizzando dei fattori di domanda e offerta; quantitativo, quando si effettuano analisi statistiche sui prezzi di mercato con lo scopo di ottenere delle previsioni. Le CTAs vengono in genere compensate attraverso delle *management fees* annuali, calcolate come percentuale sul capitale gestito, e delle *incentives fees*, calcolate come percentuale dei profitti realizzati. Un ulteriore aspetto caratteristico delle CTAs è che sono regolate dal governo degli Stati Uniti attraverso la registrazione nella Commodity Futures Trading Commission (CFTC) e l'iscrizione nella National Futures Association (NFA). In certi casi, il termine CTA potrebbe anche essere applicato agli *investment advisors* di Hedge funds, Exchange-traded funds e fondi comuni d'investimento.

output di differenti modelli (tre per gli Eurex VStoxx Futures e quattro per i CBOE Vix Futures).

- *DH Styx Program*: tale programma, lanciato nel settembre del 2017, mira ad ottenere rendimenti assoluti investendo in un paniere di FX futures scambiati nel CME (AUD/USD, CAD/USD, EUR/USD, GBP/USD, MXP/USD, NZD/USD, SWF/USD, JPY/USD[31]) e nell'ICE US (Dollar Index). Più in dettaglio, da un lato utilizza strategie di trading *contrarian* (con *time-frame* di 5 minuti) con il fine di generare profitti in quei mercati dove i prezzi oscillano su e giù senza una specifica direzione; dall'altro lato sfrutta le dinamiche di mercato di lungo periodo mediante delle strategie di trading *trend-following*, con *time-frames* di 30-60 e 120 minuti. Il programma è automatizzato dalla generazione dell'ordine fino all'esecuzione. Tale caratteristica rimuove tutti i rischi operativi connessi con l'intervento umano.

Il problema di ottimizzazione proposto riguarda proprio quest'ultimo programma. Tuttavia, prima di una sua formulazione dettagliata, si rende necessario fornire una breve introduzione sulle caratteristiche ed il funzionamento dei contratti futures.

4.2 I futures

Un future è un contratto derivato[32] stipulato tra due individui con il fine di acquistare o vendere un determinato bene (che può essere finanziario o fisico), detto sottostante, ad una certa data futura ed a un prezzo predeterminato. Tali strumenti sono definiti standardizzati in quanto scambiati in un mercato regolamentato. Questo significa che ogni caratteristica del contratto è fissata in anticipo: la tipologia del bene, la data di scadenza, la quantità, il numero massimo di contratti che è possibile scambiare in una giornata, i tick (ovvero le variazioni minime di prezzo ammesse), e così via. Inoltre il rischio di insolvenza della controparte è annullato grazie alla presenza della Cassa di Compensazione e

[31] In Appendice B la legenda delle sigle dei cambi.
[32] Sono strumenti complessi il cui prezzo si basa sul valore di mercato di un determinato bene. I principali derivati sono i futures, le opzioni e gli swap.

Garanzia (Clearing House), che funge da garante sul buon fine di ogni transazione. Diversamente dalle azioni, è una tipologia di contratto simmetrica poiché entrambe le parti sono obbligate a compiere una prestazione a scadenza. Colui che si impegna ad acquistare il sottostante ad una certa data futura assume una posizione long, mentre, dall'altro lato, il contraente che si impegna a vendere il bene assume una posizione short[33]. Dunque, l'apertura di un nuovo contratto necessita contemporaneamente di un compratore (buyer), che confida in un aumento del prezzo del sottostante, e un venditore (writer), che prevede una diminuzione del prezzo dello stesso; un future non può esistere senza la presenza di due controparti contrarie ed attive. Di conseguenza, i due contraenti hanno una visione opposta in merito alla futura evoluzione del prezzo del sottostante, e solamente uno dei due chiuderà la propria posizione con un profitto. Come tutti i derivati, i futures sono contratti fungibili, ovvero possono essere scambiati in pochi secondi da un investitore ad un altro attraverso delle semplici transazioni elettroniche. Da questo ne consegue che il buyer e il writer di uno stesso future non sono legati tra loro da un rapporto indivisibile: se, ad esempio, l'individuo A, dopo aver stipulato un future con l'individuo B, vuole uscire dalla propria posizione short, gli basta trovare un individuo C che intenda vendergli un future, affinché A possa acquistarlo e chiudere la sua posizione corta. Dopo questa operazione A si ritrova flat[34], mentre C diventa la nuova controparte di B. In altre parole se un investitore vuole chiudere la propria posizione gli basta stipulare un nuovo contratto di segno opposto rispetto al primo: questo è possibile poiché tutti i futures sono quotati in un mercato, quindi hanno un loro book dove chiunque può affacciarsi ed eseguire acquisti e vendite istantaneamente.

In genere, tali strumenti possono essere impiegati per due scopi alternativi:

- Hedging, ovvero quei casi in cui un investitore intende coprirsi dal rischio di perdere soldi a causa della variazione negativa del valore di un determinato bene. Ad esempio, un agricoltore che assume una posizione corta su un future sul grano per coprirsi dal rischio che il prezzo di questo crolli nei mesi

[33] A differenza delle azioni, la vendita allo scoperto di un future è un'operazione immediata poiché non richiede né il prestito né il pagamento di commissioni dedicate.
[34] Un investitore è in posizione flat quando non ha alcun contratto aperto.

seguenti. In genere, chi segue una strategia di questo tipo, mantiene il contratto fino alla scadenza.

- Speculazione, ossia quando il future viene utilizzato solamente con l'obiettivo di speculare sull'andamento del prezzo di un certo bene nel breve termine, senza acquistare direttamente il sottostante e, quindi, con una notevole leva finanziaria.

In quei casi in cui il sottostante è trasferibile o fisico (come le valute, i metalli, i cereali, le azioni, gli etf) può effettivamente realizzarsi lo scambio del bene alla data di consegna, tuttavia, nella maggior parte dei casi, il regolamento a scadenza è solo monetario. Inoltre, gli investitori privati non possono portare a scadenza posizioni su contratti che comportano il trasferimento fisico di merci; quando necessario, tali posizioni vengono chiuse d'ufficio dai broker.

Per quanto riguarda i futures sugli indici azionari e sulle valute, le quotazioni sono espresse rispettivamente in punti indice e tassi di cambio. Per determinare il valore monetario effettivo di tali contratti si utilizza la specifica contrattuale detta point value, che esprime quanto vale un punto dell'indice o del tasso di cambio. Ad esempio, se il future EUR/USD quota 1.18 con point value di 125000€, allora un singolo contratto varrà 147500€. Al momento della stipula del contratto le parti non versano l'intero importo del valore dello stesso, ma solo una piccola parte che è definita come margine iniziale. Quest'ultimo è un importo monetario che viene richiesto dalla Cassa di Compensazione e Garanzia (CCG), nella fase di apertura di un'operazione, sia al buyer che al writer, con il fine di coprire il rischio legato alla posizione stessa. Solitamente, e in condizioni di mercato "normali", il margine richiesto è circa del 10%, questo spiega perché tali strumenti presentano un'elevata leva intrinseca e, di conseguenza, il motivo per cui sono molto apprezzati dai traders di breve termine. Il margine iniziale è calcolato come la massima perdita giornaliera attesa sul sottostante. Tale stima esamina gli effetti che avrebbero dei forti aumenti o crolli di prezzo sul portafoglio di uno specifico investitore.

Nel periodo seguente alla stipula del contratto, il margine iniziale non rimane costante, ma può variare in seguito al verificarsi di due circostanze: la prima deriva dal fatto che tale margine è una percentuale sul valore del contratto, quindi una variazione di quest'ultimo comporterebbe un adeguamento del primo; la seconda

consiste nell'aumento del rischio connesso con il prezzo del sottostante (calcolato in automatico dai sistemi di controllo della borsa), che causerebbe un aumento della percentuale del margine. In altri termini, il valore del margine iniziale può variare sia a causa dell'andamento del prezzo del sottostante, sia in funzione dell'evoluzione del rischio del mercato. Tali adeguamenti si applicano alla fine di ogni giornata operativa. Il vero profitto (o perdita) di un future deriva dai margini di variazione giornalieri: ovvero i crediti o debiti regolati tutti i giorni, con lo scopo di non intaccare il margine iniziale. A titolo esemplificativo si veda la Tabella 1, nella quale è riportato l'andamento di una posizione long su un future E-mini S&P 500 ipotizzando un margine iniziale fisso pari al 10%, un capitale iniziale di 14000$ e un point value di 50$: i valori all'interno della terza colonna, ovvero i margini iniziali, sono pari al 10% del prodotto tra i punti indice e il point value; l'adeguamento del margine è pari alla differenza tra il margine iniziale del giorno precedente e il margine iniziale attuale; il margine di variazione corrisponde al prodotto tra il point value e il valore risultante dalla differenza tra i punti indice attuali e quelli del giorno precedente; il capitale disponibile è calcolato come somma tra l'adeguamento del margine, il margine di variazione e il capitale disponibile il giorno precedente; il capitale totale sul conto è pari alla somma tra il capitale totale disponibile il giorno precedente ed il margine di variazione.

Data	Punti indice	Margine iniziale ($)	Adeguamento margine ($)	Margine variazione ($)	Capitale disponibile ($)	Capitale totale sul conto ($)
09/07/2018	2769.5	13847.5	-	-	152.5	14000
09/07/2018	2787	13935	-87.5	875	940	14875
10/07/2018	2797.2	13986	-51	510	1399	15385
11/07/2018	2774	13870	116	-1160	355	14225
12/07/2018	2798	13990	-120	1200	1435	15425

Tabella 1: esempio di un'operazione long sul future E-mini S&P 500.

Tutto questo sistema di regolamenti giornalieri di profitti e perdite è detto *marking to market*. Nel momento in cui un investitore non possedesse sul proprio conto una quantità di denaro sufficiente a coprire il margine di variazione negativo, il broker

gli comunicherebbe che può fare solo due cose: o traferire (con effetto immediato) dei fondi da altri conti, oppure chiudere la propria posizione il giorno seguente.
In genere il prezzo di un future è diverso dal prezzo del sottostante sul quale è iscritto. Questo può verificarsi a causa delle dinamiche della domanda e dell'offerta, oppure in relazione alle caratteristiche del bene stesso; si realizza sempre la coincidenza tra i due prezzi solamente alla data di scadenza del contratto. In particolare, il prezzo di un future può essere determinato sulla base di ragionamenti matematici che mirano ad eliminare le possibili opportunità di arbitraggio. Dato che nel presente lavoro si utilizzeranno esclusivamente future sulle valute, la seguente formula esplicita il calcolo del *fair value* all'istante *t* di tale tipologia di contratti:

$$Fut_t = S_t * e^{(R_i - R_f) * (T-t)},$$

dove:

- R_i e R_f rappresentano rispettivamente il tasso di interesse nazionale e quello estero;
- S_t è il tasso di cambio spot all'epoca *t* tra la valuta nazionale ed estera;
- T è l'epoca di scadenza del contratto.

In questo caso il prezzo del future deve essere tale da neutralizzare la possibilità di guadagnare profitti extra spostando denaro da una valuta all'altra con il fine di sfruttare gli interessi più elevanti in un paese o nell'altro (di conseguenza rendendo uguali i due investimenti). Se, ad esempio, i bond europei avessero un tasso di interesse inferiore rispetto a quello dei bond americani, un investitore razionale europeo potrebbe convertire euro in dollari per comprare bond americani. Inoltre, con lo scopo di proteggersi dal rischio di cambio, comprerebbe dei future che gli consentirebbero sia di bloccare il valore del tasso di cambio sia di rimandare ad una data futura l'impiego del capitale pari al controvalore dei contratti. È evidente che se il prezzo dei futures fosse uguale a quello spot, l'investitore potrebbe beneficiare di un rendimento più elevato a costo zero. Se, all'opposto, il prezzo dei futures fosse superiore rispetto a quello spot, il beneficio del maggior tasso di interesse americano sarebbe annullato dai costi della copertura. Questo è il motivo per cui il prezzo del future (in questo caso) deve essere superiore a quello spot. A titolo esemplificativo si veda la Figura 16, la quale mostra gli andamenti congiunti di tali

prezzi, considerando come sottostante il tasso di cambio euro/dollaro.

Si noti che la formula definita precedentemente non considera le componenti connesse con le dinamiche della domanda e dell'offerta: è una semplice formula matematica legata al tasso di cambio spot, al tasso d'interesse nazionale, al tasso d'interesse estero e alla data di scadenza del contratto. Quindi non si ipotizzano variazioni di prezzo del future in seguito a temporanei sbilanci tra venditori e acquirenti. Tuttavia, il prezzo di mercato non può discostarsi eccessivamente dal *fair value* altrimenti verrebbe immediatamente "attaccato" dagli arbitraggisti, che lo riporterebbero immediatamente al suo valore di equilibrio.

In quei casi in cui il future viene mantenuto fino alla scadenza, la consegna si effettua al cosiddetto prezzo di *settlement* (o di liquidazione), ossia l'apertura del terzo venerdì del mese di scadenza: a questo punto, la controparte che aveva una posizione long diventa il compratore del bene sottostante, mentre il contraente con la posizione short diventa il venditore dello stesso. Dato che l'acquisto avviene al prezzo di *settlement*, e quindi per l'intero valore, una conseguenza fondamentale derivante dalla scadenza del contratto è il venir meno dell'effetto leva. Questo è il motivo per cui gli speculatori non tengono mai le posizioni aperte fino al giorno di consegna.

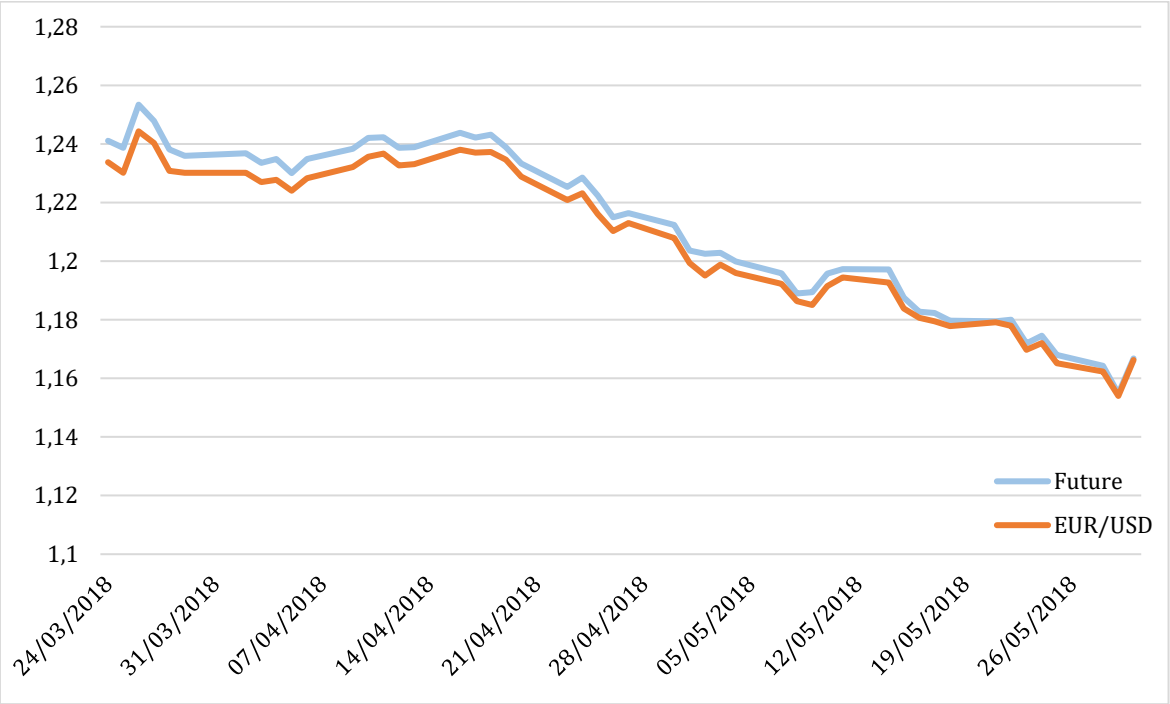


Figura 16: serie storiche del tasso di cambio EUR/USD e del valore del future iscritto sullo stesso.

Nel caso in cui un trader non avesse intenzione di chiudere una posizione long o short su una tipologia di future il cui contratto è vicino alla data di scadenza[35], questo può effettuare l'operazione di *rollover*, ovvero chiudere la posizione in essere e riaprirne una corrispondente su un contratto con scadenza più lontana. In tal modo si genera un *roll yield*[36] che può essere negativo o positivo.

4.3 Presentazione del modello di selezione

Il problema di selezione proposto da DeltaHedge sorge dall'obiettivo di determinare l'allocazione ottima di 2 milioni di euro (in leva) all'interno di Styx, ovvero il più recente *managed futures program* lanciato dalla società. Come accennato nel primo paragrafo, Styx investe in un paniere di futures sulle valute (AUD/USD, CAD/USD, EUR/USD, GBP/USD, MXP/USD, NZD/USD, SWF/USD, JPY/USD, Dollar Index) mediante due strategie di trading alternative. Più dettagliatamente, a livello macro, tale programma può essere suddiviso in due sottoprogrammi:

- il primo specula su ogni future componente il paniere attraverso una strategia di trading *contrarian* con *time-frame* di 5 minuti. Come si sarà intuito dal nome, questa strategia cerca di ottenere profitti andando controcorrente, ovvero entrare long quando tutti vendono ed aprire una posizione short quando tutti acquistano[37]. Quindi è una situazione esattamente opposta ai sistemi di trend following.
- Il secondo specula su ogni future componente il paniere mediante una strategia *trend-following* con *time-frame* che può essere di 30, 60 o 120 minuti. Questa tipologia di trading system mira ad ottenere guadagni individuando e seguendo il trend del prezzo di un determinato future.

[35] In genere questo si verifica quando la strategia di trading, sulla quale si basa l'operatività del trader, non produce un segnale di uscita da una determinata posizione (aperta precedentemente), la cui scadenza è sempre più vicina.
[36] In particolare, nel caso di posizioni long, il *roll yield* sarà negativo in presenza di contango (ovvero quando le scadenze più lontane quotano prezzi più elevati di quelle vicine) e sarà positivo in presenza di backwardation (situazioni atipiche nelle quali il prezzo dei futures sulle scadenze vicine supera quello sulle scadenze lontane).
[37] Tali strategie non condividono il famoso detto "*The trend is your friend*".

Rispetto al caso precedente, i livelli delle *stop-loss* sono in genere più ampi, in modo da sfruttare fino in fondo eventuali tendenze di lungo periodo.

Il modello di allocazione oggetto del presente lavoro riguarda solamente il primo dei due sottoprogrammi appena definiti. In particolare, la richiesta di DeltaHedge consiste nel determinare, dato un capitale iniziale, il numero di contratti ottimale (da un punto di vista rischio-rendimento) da associare ad ogni strategia di trading contrarian applicata a ciascuno dei nove FX futures componenti il paniere. Questo significa che deve essere individuato il numero ottimo di contratti mediante il quale ogni singola strategia deve operare. Se, ad esempio, si utilizzasse un'allocazione equi-pesata, attribuendo un contratto ad ogni strategia, ne conseguirebbe che nel momento in cui uno, dei nove sistemi di trading, dovesse aprire una posizione long o short su una determinata tipologia di future, questo lo farebbe su un singolo contratto.
Un'ulteriore richiesta di DeltaHedge, che si traduce in un primo vincolo, riguarda il livello di capitale che deve essere allocato all'intero sotto-programma. Dato che tutte le operazioni sono eseguite utilizzando futures, non si verifica un esborso dell'intero valore dei differenti contratti al momento dell'acquisto, perciò non sarebbe propriamente esatto determinare l'allocazione ottima sulla base di tutto l'importo monetario presente nel conto. Entrando più nel dettaglio, il primo vincolo è definito in termini di leva, la quale (seguendo le indicazioni di DeltaHedge) deve rispettare un range delimitato dai due valori 0.95-1.25, su un conto iniziale di 2000000 di dollari. Partendo dalla Tabella 2, è possibile calcolare il valore della leva di una particolare allocazione semplicemente seguendo due passi: prima si sommano tutti gli importi numerici ottenuti dal prodotto di ogni valore dell'ultima riga della tabella per la rispettiva quantità di contratti attribuiti; poi si divide la somma così ottenuta per l'importo monetario presente nel conto (20000000€).

Contratto	**6A**	**6B**	**6C**	**6E**	**6J**	**6M**	**6N**	**6S**	**DX**
Point value	100000	62500	100000	125000	125000	500000	100000	125000	1000
Tasso corrente	0.7394	1.3193	0.7586	1.1718	0.8874	0.05245	0.6794	1.0081	94.155
Valore totale ($)	73940	82456	75860	146475	110925	26225	67940	126012	94155

Tabella 2: point value, tasso di cambio attuale e valore di ogni contratto utilizzato in Styx.

Un secondo vincolo, ricavabile dalle indicazioni fornite da DeltaHedge, è connesso con la diversificazione del portafoglio ottenuto. In particolare, per evitare un'eccessiva concentrazione dei capitali a disposizione, il numero totale di strategie, sulle quali il programma investe, deve essere almeno pari a 7: quindi il modello di selezione complessivo deve considerare solamente quei portafogli composti da 7 o più strategie.

Dal momento che le strategie di trading utilizzano futures e non azioni, anche il concetto di liquidità del portafoglio diventa di fondamentale importanza, e non può non essere considerato nella determinazione dell'allocazione ottimale. Il grado di liquidità di un contratto può essere definito come una misura di rischio dello stesso in quanto rappresenta la semplicità o la complessità con cui è possibile scambiare lotti di futures di certe dimensioni, senza influenzare l'andamento del prezzo. Un esempio può rendere più chiaro tale concetto: si pensi al caso in cui il prezzo del future raggiunga un livello per il quale la strategia di trading mandi un ordine di vendita di 10 contratti. Se non ci fossero abbastanza compratori di quella quantità a quello specifico prezzo, l'investitore avrebbe solamente due alternative: la prima sarebbe quella di aspettare e sperare che altri acquirenti si facciano avanti (affrontando il rischio di essere superato da altri venditori disposti a vendere a prezzi più bassi); la seconda sarebbe quella di "scalare il book", ossia piazzare gli ordini non eseguiti sui rimanenti livelli di prezzo del book[38]. In quest'ultimo caso si genera uno *slippage* negativo, che può essere definito come il costo opportunità derivante dalla differenza tra il prezzo al quale si voleva eseguire l'ordine ed il prezzo al quale è stato effettivamente eseguito.

Le strategie di trading contrarian, che caratterizzano il programma Styx, sono tutte di tipo intra-day, di conseguenza lo *slippage* negativo cumulato[39] può diventare consistente con il passare del tempo. Esistono diversi metodi per individuare la liquidità di un determinato contratto. Molti traders osservano il cosiddetto spread denaro-lettera, ovvero la distanza tra il prezzo del miglior offerente e quello del miglior acquirente; un valore elevato di tale indicatore è un segnale di scarsa

[38] Con il termine book si intende il prospetto dove sono esposte, per un determinato strumento finanziario, le proposte di acquisto e di vendita ordinate sulla base del livello di prezzo (crescente se in vendita, decrescente se in acquisto).

[39] Ovvero la somma di tutti gli *slippage* giornalieri.

liquidità di quel particolare contratto. Un secondo metodo potrebbe essere quello di osservare, per un determinato future, il numero totale di posizioni aperte in uno specifico istante temporale (open interest). Come si può notare dalla Tabella 3, la quale mostra gli open interest dei differenti contratti scambiati dal programma Styx, il numero totale di posizioni aperte sui futures 6E è nettamente superiore rispetto a quello di altre tipologie di contratti come il 6N. Di conseguenza, anche se la strategia sul 6N fosse preferibile a quella sul 6E, bisognerebbe prestare attenzione nell'attribuire troppi contratti alla prima, poiché si potrebbe incorrere in costi elevati derivanti dalla sua scarsa liquidità.

Contratto	**6A**	**6B**	**6C**	**6E**	**6J**	**6M**	**6N**	**6S**	**DX**
Open interest	134958	189540	151448	480928	218326	163106	53344	92174	50405

Tabella 3: open interest dei 9 contratti utilizzati in Styx (scadenza settembre 2018 in data 19/07/2018). Fonte: CME group e ICE.

Ovviamente tale problema è tanto più evidente quanto maggiori sono le quantità di contratti scambiate in ogni strategia. Dato che il conto iniziale di Styx è "solo" di 2 milioni di dollari, è evidente che i costi connessi con lo *slippage* negativo sono trascurabili. Tuttavia, nel momento in cui tale importo aumentasse, si renderebbe indispensabile includere un vincolo che garantisca il rispetto di un determinato livello di liquidità del portafoglio.

Riassumendo quanto detto finora, lo scopo del presente lavoro è quello di risolvere un reale problema di ottimizzazione vincolata, nel quale la funzione obiettivo consiste nella minimizzazione di una misura di rischio two-sided (presentata nel capitolo 2), mentre l'insieme dei vincoli è definito sulla base delle richieste avanzate da DeltaHedge.

Nel seguito si riportano, da un punto di vista quantitativo, tutti i vincoli applicati al modello.

- *Vincolo sul profitto*

  Tale vincolo definisce il profitto medio minimo che l'allocazione selezionata deve soddisfare. Considerando *n* futures, con profitto atteso dell'*i-esimo* contratto pari a $\hat{p}_i$, il vincolo sul profitto può essere così definito:

$$\sum_{i=1}^{n} x_i \hat{p}_i \geq \pi.$$

- *Vincolo sulla leva*

  Questo vincolo impone che il valore della leva dell'allocazione selezionata non si discosti da un valore massimo (L_{max}) ed un valore minimo (L_{min}), ossia:

$$L_{min} \leq \frac{\sum_{i=1}^{n} V_i * x_i}{C_T} \leq L_{max},$$

  dove V_i è il valore di un singolo contratto *i*, mentre C_T è l'importo monetario totale presente nel conto dell'investitore.

- *Vincolo di cardinalità*

  Questo vincolo è introdotto per aumentare la diversificazione dell'allocazione selezionata, imponendo un numero minimo (k_d) di contratti da selezionare:

$$\sum_{i=1}^{n} z_i \geq k_d,$$

  dove z_i è un variabile binaria che assume valore 0 se il titolo *i-esimo* non è stato selezionato, altrimenti assume valore 1.

- *Vincolo sulla non negatività del numero di contratti*

  Dato che ogni allocazione ottenuta deve indicare il numero di contratti (x_i) mediante il quale ogni singola strategia deve operare, è evidente che non ha senso un x_i negativo. Questo è il motivo per cui viene inserito il seguente vincolo:

$$x_i \geq 0 \; \forall i \in R.$$

- *Vincolo d'integralità*

  Una possibile formulazione della variabile $z_i \in \{0,1\}$ è la seguente (Corazza *et al.*, 2013):

$$z_i(z_i - 1) = 0 \; \forall i \in R.$$

- *Vincolo sulle percentuali d'investimento*

  Questo vincolo mira a contenere la percentuale di capitale attribuita ad ogni strategia. Data una frazione massima u_i e minima d_i d'investimento possibile sulla strategia *i-esima*, tale vincolo può essere formulato nel seguente modo:

$$z_i d_i \leq \frac{v_i * x_i}{C_P} \leq z_i u_i \; \forall i \in R$$

$$con \; 0 \leq d_i \leq u_i \leq 1 \text{ e } u_i \geq \frac{1}{n},$$

  dove $C_P = \sum_{i=1}^{n} V_i * x_i$, ovvero il capitale totale richiesto da quella specifica allocazione. Inoltre, affinché tale vincolo sia compatibile con quello sulla cardinalità, deve valere la seguente disuguaglianza:

$$d_i \leq \frac{1}{k_d}.$$

  Quindi, se un investitore volesse inserire almeno 7 contratti (k_d) nel portafoglio finale, la percentuale minima di ciascuno non può essere superiore a circa il 14% (d_i).

A questo punto è possibile presentare, in termini matematici, il problema di selezione complessivo, la cui risoluzione sarà oggetto dei successivi paragrafi:

$$\min_{X,Z} \; \frac{a}{T}[\sum_{t=1}^{T}(\sum_{i=1}^{n}(p_{it} - \hat{p}_i)x_i)^+] + (1-a)\left\{\frac{1}{T}\sum_{t=1}^{T}[(\sum_{i=1}^{n}(p_{it} - \hat{p}_i)x_i)^-]^p\right\}^{\frac{1}{p}} - \sum_{i=1}^{n} x_i * \hat{p}_i$$

$$s.t. \begin{cases} \sum_{i=1}^{n} x_i \hat{p}_i \geq \pi \\ L_{min} \leq \dfrac{\sum_{i=1}^{n} V_i * x_i}{C_T} \leq L_{max} \\ \sum_{i=1}^{n} z_i \geq k_d \\ x_i \geq 0 \; \forall i \in R \\ z_i(z_i - 1) = 0 \; \forall i \in R \\ z_i d_i \leq \dfrac{v_i * x_i}{C_P} \leq z_i u_i \; \forall i \in R \end{cases}.$$

L'allocazione ottima, ricavabile dalla risoluzione di tale modello di selezione,

potrebbe essere composta da un numero di contratti, attribuiti ad ogni strategia, non intero. Ad esempio, si potrebbe ottenere un portafoglio che attribuisce 5.2 contratti alla strategia di trading contrarian sui futures 6E; ovviamente questo non è possibile nella realtà, in quanto si possono acquistare e vendere solamente quantità intere di una certa tipologia di futures. Nonostante questo ulteriore vincolo non sia stato inserito nel modello di selezione complessivo, verrà comunque considerato seguendo quanto fatto da Laskari *et al.* (2002): nel momento in cui le due metaeuristiche determinano una nuova soluzione, i contratti attribuiti ad ogni strategia vengono approssimati al numero intero più vicino.

Come accennato precedentemente, l'insieme di vincoli a variabili intere che caratterizzano questo problema di minimizzazione, lo rendono sia non differenziabile sia non lineare. Da questo ne deriva la necessità di sfruttare dei metodi approssimati per la sua risoluzione, ovvero le metaeuristiche.

4.4 Riformulazione del modello di selezione

La PSO e la BFO sono due metaeuristiche che permettono di risolvere problemi di ottimizzazione non vincolata. Di conseguenza, prima di utilizzarle per la risoluzione del problema di minimizzazione presentato precedentemente, si rende necessario effettuare un ulteriore passaggio, che si sostanzia in una riformulazione del modello vincolato in una forma non vincolata. Tra tutti i metodi proposti in letteratura che perseguono tale scopo[40], la scelta è ricaduta sul metodo delle penalità esatte. In linea di massima, tale metodo consiste nell'aggiungere alla funzione obiettivo la somma di tutte le violazioni dei vincoli moltiplicate per uno specifico parametro ε. In questo modo, oltre a minimizzare il valore associato alla misura di rischio, si minimizza il livello di tutte le violazioni dei vincoli. In termini formali, il criterio delle penalità esatte equivale a riscrivere un generico problema di ottimizzazione vincolata del seguente tipo (Luenberger, 1984):

$$\min f(x)$$

[40] Tutte le metodologie possono essere suddivise in 4 categorie (Hu ed Eberhart, 2002): metodi numerici che cercano di mantenere solamente le soluzioni ammissibili; criteri che distinguono tra soluzioni ammissibili e non ammissibili; metodi che si fondano su funzioni di penalità; metodi ibridi.

$$s.t.\begin{cases} h(x) = 0 \\ g(x) \leq 0 \end{cases}$$

dove:

- $f(x)$ è una funzione continua in R^n;
- $h(x)$ è un vettore composto da m vincoli espressi in forma di uguaglianza;
- $g(x)$ è un vettore composto da p vincoli definiti in forma di disuguaglianza;

come un problema di ottimizzazione non vincolata formulato nel seguente modo:

$$\min f(x) + \varepsilon P(x)$$
$$con\ P(x) \geq 0\ \forall x \in R^n$$

dove:

- ε è una costante positiva detta fattore di penalità;
- $P(x)$ è una funzione continua in R^n.

Il ruolo del fattore di penalità è quello di aumentare il valore numerico di ogni singola violazione. La corretta scelta di tale parametro è di fondamentale importanza, poiché garantisce la corrispondenza tra la soluzione del problema vincolato e quello riformulato.
In particolare, formulando la funzione di penalità $P(x)$ nel seguente modo:

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

se un generico punto x^* soddisfa le condizioni sufficienti del secondo ordine per essere un punto di minimo locale del problema vincolato, lo sarà anche per il problema non vincolato quando

$$\varepsilon > max\{|\omega_i|, \mu_j : i = 1, \dots, m, j = 1, \dots, p\},$$

dove ω e μ sono due vettori composti da moltiplicatori di Lagrange relativi, rispettivamente, agli m vincoli compresi in $h(x) = 0$ ed ai p vincoli all'interno di $g(x) \leq 0$.
Come si può notare, questo teorema da un lato presenta le condizioni sufficienti affinché le soluzioni delle due tipologie di problemi coincidano, però, dall'altro lato, non fornisce alcuna indicazione su come determinare il valore di ε.

Per quanto riguarda l'applicazione del metodo delle penalità esatte al problema di ottimizzazione vincolata presentato nel precedente paragrafo, il primo passo consiste nel riscrivere tutti i vincoli in termini di violazioni:

$$\sum_{i=1}^{n} x_i \hat{p}_i \geq \pi \rightarrow \max\{0, \pi - \hat{p}\}$$

$$L_{min} \leq \frac{\sum_{i=1}^{n} V_i * x_i}{C_T} \rightarrow max\left\{0, L_{min} - \frac{\sum_{i=1}^{n} V_i * x_i}{C_T}\right\}$$

$$\frac{\sum_{i=1}^{n} V_i * x_i}{C_T} \leq L_{max} \rightarrow max\left\{0, \frac{\sum_{i=1}^{n} V_i * x_i}{C_T} - L_{max}\right\}$$

$$\sum_{i=1}^{n} z_i \geq k_d \rightarrow \max\left\{0, k_d - \sum_{i=1}^{n} z_i\right\}$$

$$x_i \geq 0 \rightarrow \sum_{i=1}^{n} \max\{0, -x_i\}$$

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

$$z_i d_i \leq \frac{v_i * x_i}{C_P} \rightarrow \sum_{i=1}^{n} \max\left\{0, z_i d_i - \frac{v_i * x_i}{C_P}\right\}$$

$$\frac{v_i * x_i}{C_P} \leq z_i u_i \rightarrow \sum_{i=1}^{n} \max\left\{0, \frac{v_i * x_i}{C_P} - z_i u_i\right\}$$

A questo punto, è possibile riformulare il modello di selezione iniziale, in forma non vincolata, nel seguente modo:

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

dove:

$$P(X, Z, \varepsilon) = p_{a,p}(\hat{P})$$
$$+ \frac{1}{\varepsilon}\left[\max\{0, \pi - \hat{p}\} + max\left\{0, L_{min} - \frac{\sum_{i=1}^{n} V_i * x_i}{C_T}\right\}\right.$$
$$+ max\left\{0, \frac{\sum_{i=1}^{n} V_i * x_i}{C_T} - L_{max}\right\} + \max\left\{0, k_d - \sum_{i=1}^{n} z_i\right\}$$
$$+ \sum_{i=1}^{n} \max\{0, -x_i\} + \sum_{i=1}^{n} |z_i(1 - z_i)| + \sum_{i=1}^{n} \max\left\{0, z_i d_i - \frac{v_i * x_i}{C_P}\right\}$$
$$\left. + \sum_{i=1}^{n} \max\left\{0, \frac{v_i * x_i}{C_P} - z_i u_i\right\}\right]$$

Tale modello può essere definito come la riformulazione non vincolata del problema di minimizzazione proposto da DeltaHedge. La somma algebrica all'interno delle parentesi quadre è nulla solamente quando tutti i vincoli sono rispettati; all'opposto, se uno di questi non fosse soddisfatto, tale somma genererebbe un valore positivo che, amplificato dal fattore $\frac{1}{\varepsilon}$, andrebbe a sommarsi alla misura di rischio. Se, ad esempio, il profitto medio ($\hat{p}$) dell'allocazione ottenuta fosse inferiore al profitto minimo (π), il primo vincolo non sarebbe rispettato. Di conseguenza, si andrebbe a sommare alla misura di rischio il prodotto tra $\frac{1}{\varepsilon}$ e $\pi - \hat{p}$. È evidente che gli algoritmi scartano più facilmente quelle soluzioni che non rispettano i vincoli, in quanto la violazione di quest'ultimi comporta un aumento notevole del valore della fitness (che deve essere minimizzata).

Nel prossimo capitolo si applicherà la PSO e la BFO al modello di selezione riformulato, con il fine di determinare una soluzione approssimata che, da un lato, sia realmente utilizzabile da parte di DeltaHedge e, dall'altro lato, sia il più possibile vicina alla soluzione ottima globale.

CAPITOLO 5

DETERMINAZIONE DELL'ALLOCAZIONE FINALE ALL'INTERNO DI STYX

In questo ultimo capitolo si applicheranno la BFO e la PSO ai dati forniti da DeltaHedge, con il fine di risolvere il modello di selezione formulato nel precedente capitolo. Dopo aver definito i valori dei parametri settati per ogni metaeuristica, si verificherà se una delle due offre delle soluzioni migliori rispetto all'altra. Infine, si utilizzerà la metaeuristica "migliore", scelta nel passo precedente, per la determinazione dell'allocazione ottima finale per il programma Styx.

5.1 Informazioni iniziali

Lo scopo di questo capitolo finale è quello di fornire una possibile risposta alle richieste avanzate da DeltaHedge, ovvero determinare il numero di contratti futures ottimale (da un punto di vista rischio-rendimento) mediante il quale ogni singola strategia di trading contrarian, componente il più generale programma Styx, deve operare. In altri termini, si deve fornire a DeltaHedge la soluzione del problema di ottimizzazione vincolata introdotto nel precedente capitolo. Tale allocazione finale sarà utilizzata direttamente da DeltaHedge per investire 2 milioni di dollari (in leva) in Styx, il quale è stato inizialmente avviato associando quattro contratti futures ad ogni strategia di trading, escludendo quella sul 6J e quella sul DX (vedi Tabella 4)[41].

Strategia	**6A**	**6B**	**6C**	**6E**	**6J**	**6M**	**6N**	**6S**	**DX**
Numero contratti	4	4	4	4	0	4	4	4	0

Tabella 4: numero di contratti futures attribuiti inizialmente ad ogni strategia componente il programma Styx.

A questo punto si può apprezzare l'importanza applicativa di tale lavoro e, di

[41] La scelta di non investire nelle strategie sul 6J e sul DX è giustificata dalle loro scarse performance *in-sample.*

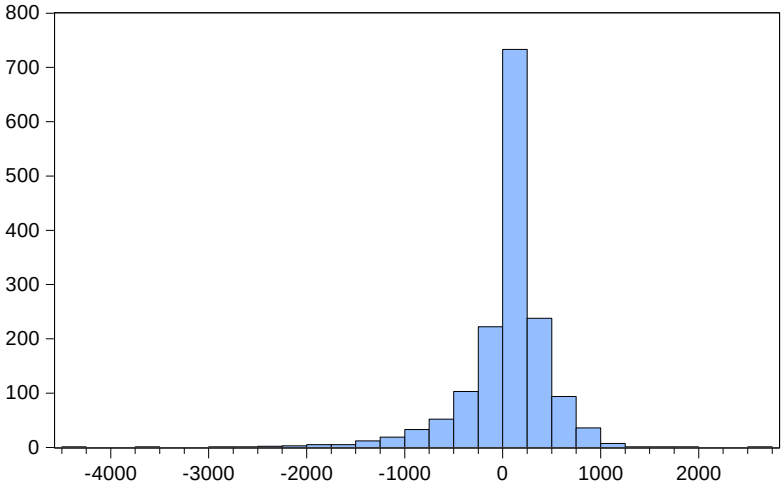
conseguenza, la necessità di raggiungere un'allocazione che non si discosti eccessivamente dalla soluzione ottima globale: la distanza tra queste due soluzioni può essere vista come un costo opportunità, che può condurre a performance anche molto differenti nel lungo periodo. Da qui si possono intuire le motivazioni sottostanti alla scelta di utilizzare due metaeuristiche, anziché una, per la ricerca della soluzione "ottima", che possono essere riassunte come un tentativo di ridurre i limiti connessi con tali tipologie di algoritmi.
Alla luce di queste premesse, si può definire tale sezione conclusiva come un processo composto dalle seguenti tre fasi:

1. verifica se entrambe le metaeuristiche rispettano le proprietà connesse con la misura di rischio two-sided;
2. applicazione della BFO e della PSO ai dati forniti da DeltaHedge con il fine di determinare l'algoritmo che offre i risultati migliori;
3. determinazione dell'allocazione finale attraverso il metodo scelto nel passo precedente.

Le analisi sono svolte utilizzando i dati forniti da DeltaHedge, che consistono nelle serie storiche dal 02/07/2012 al 10/07/2018 dei profitti giornalieri (espressi in dollari) di ognuna delle nove strategie di trading contrarian che compongono Styx[42]. Tali profitti, oltre ad essere al netto delle commissioni e al lordo[43] dello *slippage*, sono stati ottenuti revisionando periodicamente i parametri delle singole strategie. Inoltre, è importante evidenziare che tutte le strategie sono settate in modo tale che non effettuino alcun *trade* in quei giorni particolari in cui si verifica un determinato evento: ad esempio, la strategia sul 6E non effettua operazioni quando la BCE svolge delle importanti conferenze stampa. Queste giornate di *no-trade* sono state comunque considerate nel campione di dati poiché, anche se il profitto è (ovviamente) nullo, sono in ogni caso dei risultati derivanti dalla tipologia di strategia. Le figure seguenti mostrano gli istogrammi e le statistiche descrittive delle

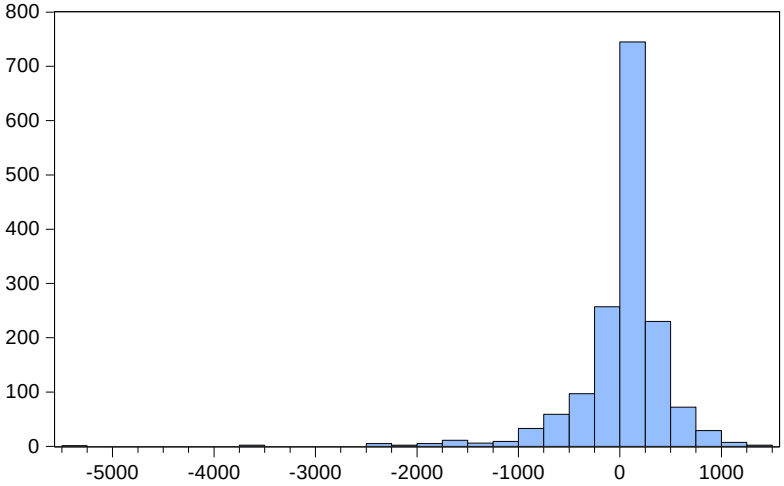
[42] Le serie storiche dei profitti di tutte le strategie sono state ottenute considerando un singolo contratto per ogni strategia.
[43] Le serie storiche fornite sono dei profitti giornalieri "virtuali", ovvero ottenuti dalla simulazione delle strategie di trading negli ultimi sei anni, quindi non è possibile conoscere lo *slippage* generato da ognuna di queste poiché tali simulazioni ipotizzano che ogni *trade* sia stato effettuato al prezzo effettivo a cui doveva essere eseguito.

nove serie storiche dei profitti giornalieri oggetto di analisi.



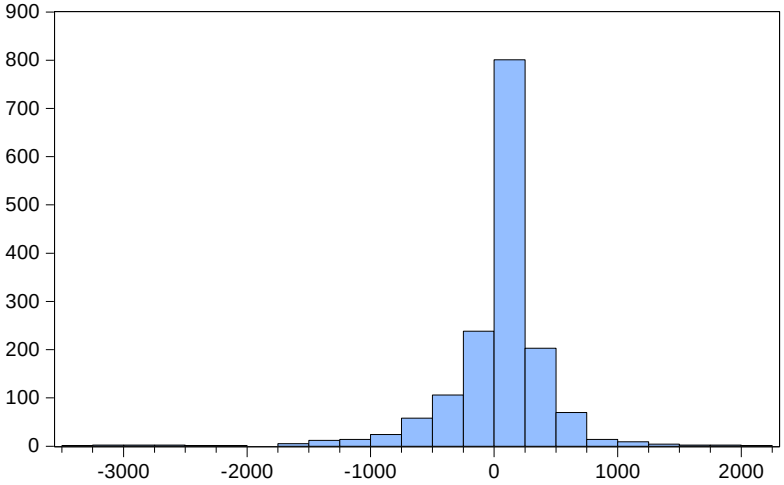
Series: _6A	
Sample 7/02/2012 7/10/2018	
Observations 1572	
Mean	22.99746
Median	7.000000
Maximum	2726.000
Minimum	-4302.000
Std. Dev.	464.8362
Skewness	-1.946705
Kurtosis	15.34676
Jarque-Bera	10977.87
Probability	0.000000

Figura 17: istogramma ed analisi descrittive dei profitti storici della strategia sul 6A.



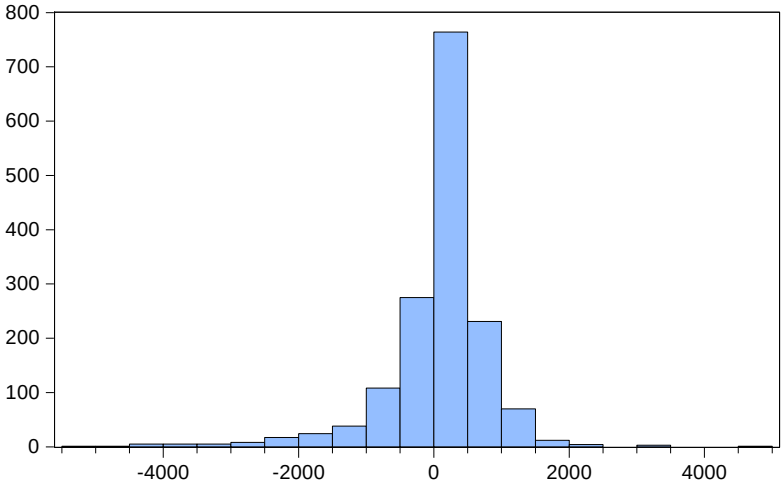
Series: _6B	
Sample 7/02/2012 7/10/2018	
Observations 1572	
Mean	7.790394
Median	8.500000
Maximum	1269.250
Minimum	-5336.000
Std. Dev.	453.5023
Skewness	-2.905861
Kurtosis	24.29933
Jarque-Bera	31927.16
Probability	0.000000

Figura 18: istogramma ed analisi descrittive dei profitti storici della strategia sul 6B.



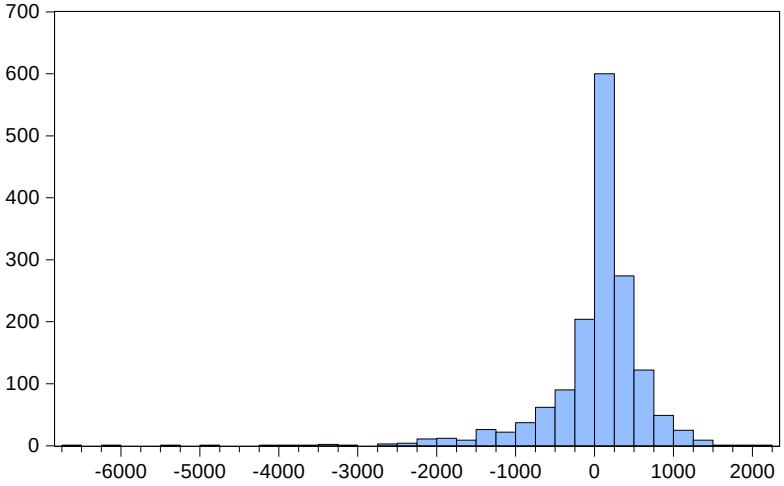
Series: _6C	
Sample 7/02/2012 7/10/2018	
Observations 1572	
Mean	12.11387
Median	0.000000
Maximum	2128.000
Minimum	-3288.000
Std. Dev.	426.0521
Skewness	-1.917166
Kurtosis	16.03840
Jarque-Bera	12097.99
Probability	0.000000

Figura 19: istogramma ed analisi descrittive dei profitti storici della strategia sul 6C.



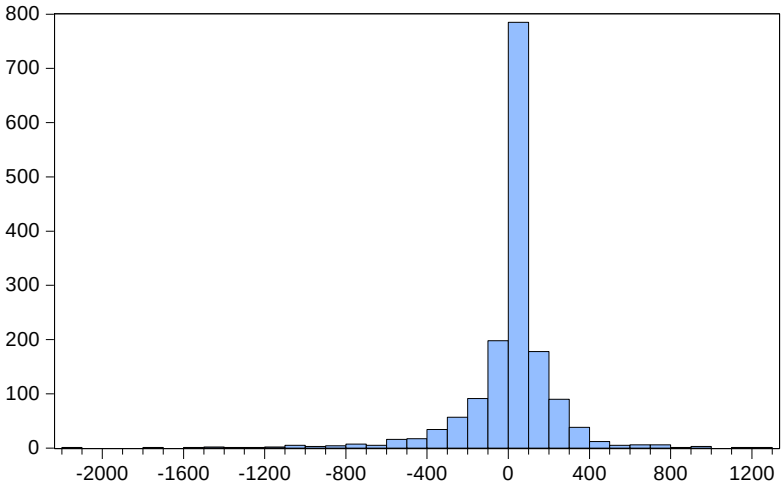
Series: _6E	
Sample 7/02/2012 7/10/2018	
Observations 1572	
Mean	41.74841
Median	54.50000
Maximum	4632.250
Minimum	-5498.500
Std. Dev.	786.8236
Skewness	-1.634658
Kurtosis	11.71619
Jarque-Bera	5676.252
Probability	0.000000

Figura 20: istogramma ed analisi descrittive dei profitti storici della strategia sul 6E.



Series: _6J	
Sample 7/02/2012 7/10/2018	
Observations 1572	
Mean	-0.685115
Median	29.75000
Maximum	2001.500
Minimum	-6574.000
Std. Dev.	666.7852
Skewness	-2.926116
Kurtosis	21.34877
Jarque-Bera	24295.65
Probability	0.000000

Figura 21: istogramma ed analisi descrittive dei profitti storici della strategia sul 6J.



Series: _6M	
Sample 7/02/2012 7/10/2018	
Observations 1572	
Mean	1.371501
Median	0.000000
Maximum	1220.000
Minimum	-2180.000
Std. Dev.	234.4497
Skewness	-1.909947
Kurtosis	17.50006
Jarque-Bera	14727.24
Probability	0.000000

Figura 22: istogramma ed analisi descrittive dei profitti storici della strategia sul 6M.

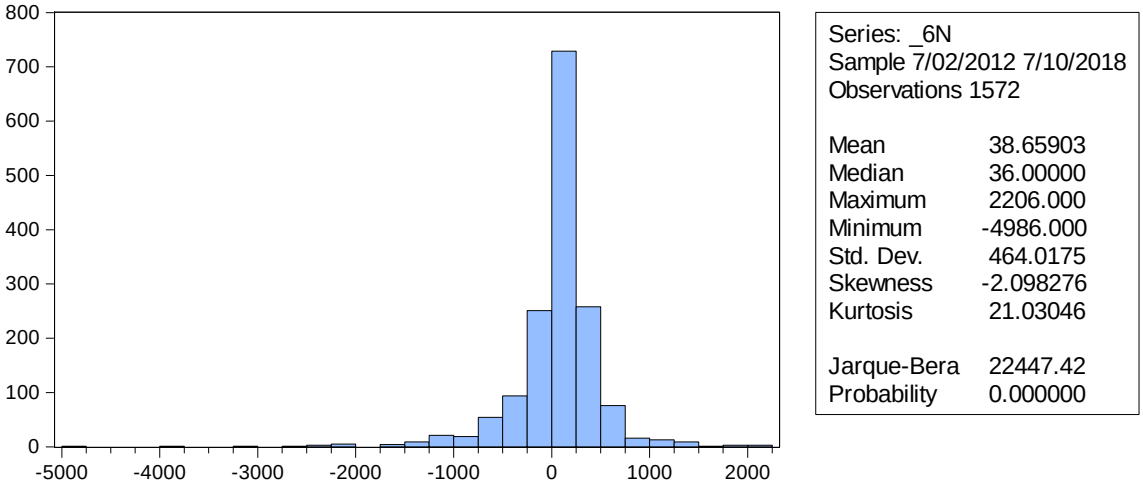


Figura 23: istogramma ed analisi descrittive dei profitti storici della strategia sul 6N.

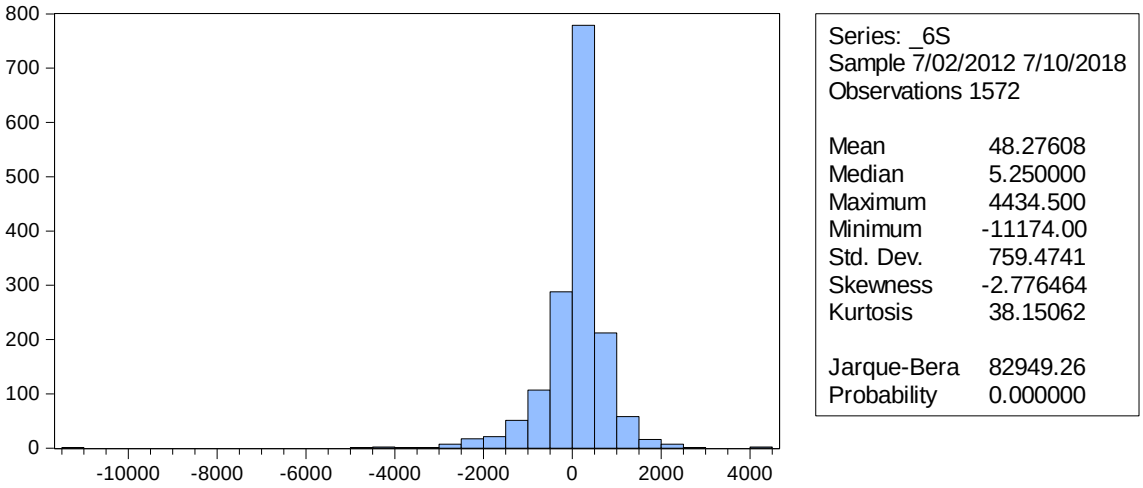


Figura 24: istogramma ed analisi descrittive dei profitti storici della strategia sul 6S.

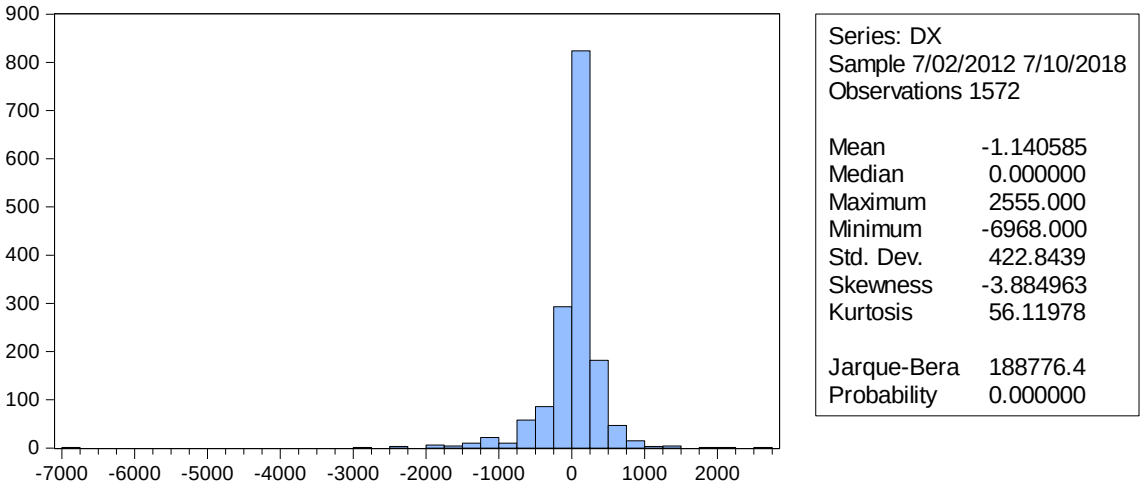


Figura 25: istogramma ed analisi descrittive dei profitti storici della strategia sul DX.

Come si può notare, le distribuzioni di tutte le serie storiche dei profitti giornalieri,

oltre a non essere normali[44], presentano una forte asimmetria negativa ed una curtosi molto elevata. Di conseguenza, se si utilizzasse il classico modello di Markowitz per la determinazione del numero ottimo di contratti per ogni strategia, si rischierebbe di sottostimare pesantemente il rischio.
Prima di proseguire con le analisi, si rende necessario presentare i parametri impostati sia per i vincoli del modello sia per le singole metaeuristiche.

5.2 Impostazione dei parametri

In questo paragrafo si intende presentare i due insiemi di parametri impostati con il fine di ottenere l'allocazione ottima. Più in dettaglio, il primo insieme contiene tutti i parametri che riguardano i vincoli del modello, i quali sono comuni ad entrambe le metaeuristiche e sono mantenuti costanti. Il secondo insieme riguarda i metodi risolutivi, ovvero quei parametri necessari per il corretto funzionamento della PSO e della BFO. Entrambe le metaeuristiche sono state processate in MATLAB 2018 mediante un HP Pavilion con processore AMD A10.[45]
Per quanto riguarda i vincoli del modello, sono stati settati (e mantenuti costanti) i seguenti valori:

- Numero minimo di strategie di trading sul quale si deve investire: $k_d = 7$;
- Livello minimo della leva: $L_{min} = 0.95$;
- Livello massimo della leva: $L_{max} = 1.25$;
- Capitale totale presente nel conto: $C_T = 2000000$;
- Percentuale di investimento minima in ogni strategia: $P_{min} = 0.01$;
- Percentuale di investimento massima in ogni strategia: $P_{max} = 0.31$.

Inoltre, seguendo le indicazioni fornite da DeltaHedge, il vincolo sul profitto obiettivo π è stato settato come:

$$\pi = \left(\sum_{i=1}^{9} \hat{p}_i * 4\right) - (\hat{p}_{6J} * 4) - (\hat{p}_{DX} * 4),$$

[44] Si veda il test di Jarque-Bera.
[45] In Appendice C sono riportati entrambi i codici utilizzati.

dove:

- $\hat{p}_{6J}$ è il profitto medio giornaliero della strategia di trading sul 6J;
- $\hat{p}_{DX}$ è il profitto medio giornaliero della strategia di trading sul DX.

In altri termini, l'allocazione ottenuta, attraverso le metaeuristiche, deve fornire un profitto medio giornaliero superiore rispetto a quello ottenuto dall'allocazione iniziale, la quale attribuisce quattro contratti su tutte le strategie ad esclusione di quella sul 6J e quella sul DX.[46]

Per quanto riguarda il valore totale di ogni tipologia di future, questo è stato ricavato utilizzando il tasso di cambio corrispondente alla data iniziale dei profitti *out-of-sample*: ad esempio, se si volesse determinare l'allocazione ottima utilizzando come campione *in-sample* i profitti ottenuti dal 10/01/2017 al 10/01/2018, per determinare il valore totale di ogni singolo future si utilizzerebbe il tasso di cambio al 10/01/2018[47].

Parametri PSO

Per quanto riguarda la prima metaeuristica, ovvero la PSO, si è deciso di utilizzare la versione con l'*Inertia Weight*. I parametri necessari per la determinazione di V_j^{k+1} sono stati settati in linea con quanto suggerito dalla letteratura:

- peso d'inerzia: $\omega = 0.7298$;
- coefficiente di accelerazione cognitivo: $c_1 = 1.49618$;
- coefficiente di accelerazione sociale: $c_2 = 1.49618$;
- numero di iterazioni: $\text{niter} = 3000$[48];
- numero di particelle: $\text{P} = 30$.

Un ulteriore parametro che deve essere impostato è il coefficiente di penalità ε, il quale assicura la coincidenza tra il problema originale in forma vincolata ed il problema riformulato con il metodo delle penalità esatte. Per una sua corretta

[46] Si veda la Tabella 4.

[47] In Appendice D sono riportati tutti i tassi di cambio utilizzati per ogni periodo *out-of-sample*.

[48] La scelta di non utilizzare un valore di iterazioni più elevato deriva, da un lato, dall'elevato tempo macchina richiesto da ogni prova del codice e, dall'altro lato, dal fatto che i valori delle fitness delle soluzioni ottenute, aumentando il numero di iterazioni, in media sono molto simili a quelli ottenuti nel caso di 3000 iterazioni.

determinazione si è deciso di effettuare cinque prove dell'algoritmo per diversi valori di ε. La Tabella 5 mostra i risultati ottenuti impostando, per quanto riguarda la misura di rischio, $a = 0.5$ e $p = 2$.

ε	Fitness normalizzata	Standard deviation
1	0.503	0.108449
0,1	0.3932	0.195563
0,01	0.3842	0.074085
0,001	0.1576	0.007157
0,0001	0.1466	0.002789
0,00001	0.1483	0.003549
0,000001	0.149	0.004447

Tabella 5: scelta del parametro ε della PSO.

Come si può notare, la media delle fitness normalizzate che ha registrato il valore più basso è stata ottenuta in corrispondenza di $\varepsilon = 0.0001$.

Parametri BFO

Diversamente da quanto visto per la PSO, la letteratura inerente alla BFO è piuttosto scarna, e non offre molte indicazioni sul corretto settaggio dei vari parametri che la caratterizzano. Di conseguenza, si è scelto di procedere effettuando varie prove dell'algoritmo, con il fine di selezionare quei valori che conducono ai risultati migliori. I parametri scelti sono i seguenti:

- numero step di chemiotassi: $N_c = 400$[49];
- numero step di avanzamento: $N_s = 4$;
- numero step di riproduzione: $N_{re} = 4$;
- numero step di *elimination-dispersal*: $N_{ed} = 4$;
- probabilità di eliminazione: $P_{ed} = 0.3$;
- ampiezza step di capovolgimento: $C_t = 0.02$;

[49] Ad una prima osservazione, 400 step di chemiotassi possono sembrare pochi. Tuttavia, considerando che 4 fasi di *elimination-dispersal* incorporano 16 step di riproduzione, in totale vengono effettuati 6400 step di chemiotassi.

- ampiezza step di avanzamento: $C_s = 0.0002$;
- numero di batteri: $b = 30$.

Come per la PSO, anche la BFO richiede la determinazione del fattore di penalità ε. A tal fine è stata effettuata la stessa procedura seguita nel caso precedente, ottenendo i risultati mostrati in Tabella 6.

ε	Fitness normalizzata	Standard deviation
1	0.701	0.062613
0,1	0.701	0
0,01	0.45741	0.093497
0,001	0.200336	0.004959
0,0001	0.196687	0.033003
0,00001	0.229132	0.024206
0,000001	0.198778	0.04201

Tabella 6: scelta del parametro ε della BFO.

Come per la PSO, la media delle fitness normalizzate che ha registrato il valore più basso è stata ottenuta in corrispondenza di $\varepsilon = 0{,}0001$.

5.3 Applicazione delle metaeuristiche

Il passo successivo all'impostazione dei parametri consiste nel verificare se il modello di selezione rispetta le due proprietà che caratterizzano la misura di rischio two-sided. In particolare, si vuole testare sia la non-crescenza di $\rho_{\alpha,p}$ rispetto al parametro α, sia la non-decrescenza di $\rho_{\alpha,p}$ rispetto al parametro p. Per svolgere queste analisi sono stati generati i seguenti sotto-campioni:

- il primo campione include tutti i profitti giornalieri dal 10/07/2017 al 10/07/2018;
- il secondo campione considera tutti i profitti giornalieri dal 07/07/2016 al 07/07/2017;

- il terzo campione comprende tutti i profitti giornalieri dal 07/07/2015 al 06/07/2016.

Iniziando dalla PSO, la Tabella 7 mostra come varia la misura di rischio $\rho_{\alpha,p}$ all'aumentare di p[50]: tali valori sono stati ottenuti con α costante posto uguale a 0.5. Come si può notare, la proprietà di monotonia di $\rho_{\alpha,p}$ è soddisfatta in tutti i sotto-campioni, in quanto all'aumentare del valore del parametro p, il valore della misura di rischio aumenta sempre.

CAMPIONE 1	***p = 1***	***p = 2***	***p = 3***
$\boldsymbol{\rho_{0.5,p}}$	1224.5	2157.8	2875.4
Strategie selezionate	9	8	9

CAMPIONE 2	***p = 1***	***p = 2***	***p = 3***
$\boldsymbol{\rho_{0.5,p}}$	956.96	2027.2	3345.3
Strategie selezionate	8	7	8

CAMPIONE 3	***p = 1***	***p = 2***	***p = 3***
$\boldsymbol{\rho_{0.5,p}}$	1206.3	2287.2	3148.4
Strategie selezionate	9	9	8

Tabella 7: verifica della non-decrescenza di $\rho_{0.5,p}$ rispetto a p (PSO).

La Tabella 8 mostra come variano i valori della misura di rischio $\rho_{\alpha,p}$ al variare del parametro α e con p costante pari a 2. Come si può vedere, anche in questo secondo caso i risultati ottenuti sono coerenti con quelli a livello teorico: in tutti i campioni il valore della misura di rischio diminuisce all'aumentare del parametro α.
Di conseguenza, entrambi i parametri possono essere utilizzati con il fine di modellizzare la propensione al rischio del singolo investitore.

CAMPIONE 1	***α = 0***	***α = 0.5***	***α = 1***
$\boldsymbol{\rho_{\alpha,2}}$	3018	2157.8	1192.7
Strategie selezionate	9	8	9

[50] Dato l'elevato tempo macchina richiesto per ogni prova, queste analisi sono state effettuate impostando un numero di iterazioni pari a 2000.

CAMPIONE 2	***α = 0***	***α = 0.5***	***α = 1***
$\boldsymbol{\rho_{\alpha,2}}$	2983.9	2027.2	935.66
Strategie selezionate	8	7	7

CAMPIONE 3	***α = 0***	***α = 0.5***	***α = 1***
$\boldsymbol{\rho_{\alpha,2}}$	3175	2287.2	1234.8
Strategie selezionate	8	9	8

Tabella 8: verifica della non-crescenza di $\rho_{\alpha,2}$ rispetto al parametro α (PSO).

Per quanto concerne la BFO, le Tabelle 9 e 10 riportano i risultati ottenuti[51]. Anche in questo caso entrambe le proprietà definite precedentemente sono rispettate in tutti i sotto-campioni.

CAMPIONE 1	***p = 1***	***p = 2***	***p = 3***
$\boldsymbol{\rho_{0.5,p}}$	1388.2	2913.2	4709.5
Strategie selezionate	7	8	7

CAMPIONE 2	***p = 1***	***p = 2***	***p = 3***
$\boldsymbol{\rho_{0.5,p}}$	1238.2	2852.9	3559.8
Strategie selezionate	8	7	7

CAMPIONE 3	***p = 1***	***p = 2***	***p = 3***
$\boldsymbol{\rho_{0.5,p}}$	1967.3	2527.9	5116.1
Strategie selezionate	7	8	7

Tabella 9: verifica della non-decrescenza di $\rho_{0.5,p}$ rispetto a p (BFO).

CAMPIONE 1	***α = 0***	***α = 0.5***	***α = 1***
$\boldsymbol{\rho_{\alpha,2}}$	3578.7	2913.2	1721.5
Strategie selezionate	7	8	7

[51] Dato l'elevato tempo macchina richiesto per ogni prova, queste analisi sono state effettuate impostando un numero di step di chemiotassi pari a 300.

CAMPIONE 2	***α = 0***	***α = 0.5***	***α = 1***
$\boldsymbol{\rho_{\alpha,2}}$	4076.3	2852.9	1463.7
Strategie selezionate	7	7	7

CAMPIONE 3	***α = 0***	***α = 0.5***	***α = 1***
$\boldsymbol{\rho_{\alpha,2}}$	4158.5	2527.9	1315
Strategie selezionate	7	8	7

Tabella 10: verifica della non-crescenza di $\rho_{\alpha,2}$ rispetto al parametro α (BFO).

Un secondo aspetto, che si può notare dall'osservazione delle precedenti tabelle, è la tendenza della PSO ad individuare allocazioni più diversificate rispetto alla BFO. Per quanto riguarda tutte le successive analisi, si è deciso di impostare $\alpha = 0.5$ e $p = 2$.

5.4 Confronto ed analisi

Questa sezione mira a determinare la metaeuristica che verrà utilizzata successivamente per il calcolo dell'allocazione ottima finale per il programma Styx. Quindi verranno confrontati i risultati ottenuti mediante la PSO e la BFO, per verificare se una offre delle soluzioni migliori rispetto all'altra. In altri termini, si vuole verificare se una metaeuristica si avvicina maggiormente alla soluzione ottima globale rispetto all'altra.
Seguendo la pratica finanziaria, si è scelto di proseguire nel seguente modo:

- si determina l'allocazione ottima utilizzando come dati *in-sample* i profitti giornalieri di un anno intero[52];
- si investe l'allocazione ottima ottenuta nel passo precedente per i successivi tre mesi, i quali rappresentano il futuro virtuale;
- si ripete il procedimento di selezione utilizzando, come dati *in-sample*, i profitti giornalieri di un altro periodo della durata di un anno (quindi, rispetto al punto uno, si sostituiscono i tre mesi meno recenti con quelli

[52] Effettuando delle prove con un campione *in-sample* di profitti giornalieri rispettivamente di due e tre anni, sono state generate delle allocazioni che hanno fornito dei risultati *out-of-sample* nettamente peggiori rispetto all'utilizzo di un campione *in-sample* di solo un anno.

utilizzati come futuro virtuale nel passo precedente), e si reinveste l'allocazione ottenuta per i successivi tre mesi;

- si ripete tutto il procedimento;
- si confrontano le due metaeuristiche sulla base delle performance ottenute su tutti i futuri virtuali.

Come si sarà intuito, tale procedimento non è altro che una simulazione sui dati storici di quello che verrà fatto a livello pratico nella realtà. Entrando più nel dettaglio, come campione di partenza si sono utilizzati i profitti giornalieri disponibili dal 01/01/2013 al 01/01/2014; quindi si è ipotizzato di essere nella sera del 01/01/2014 pronti ad investire 2 milioni di dollari, per i successivi tre mesi, nelle nove strategie di trading che compongono Styx utilizzando un'allocazione ottima individuata attraverso le metaeuristiche. Le Tabelle 11 e 12 mostrano i risultati ottenuti rispettivamente con la PSO e la BFO[53].

Prova	***1***	***2***	***3***
Fitness	1691.1	1709.4	1732.4
Misura di rischio out-of-sample	926.6972	854.4164	842.7894
Misura di rischio in-sample	1691.1	1709.4	1732.4
Profitti out-of-sample	1342	1492.5	1408.3
Profitti in-sample	1274	1285.1	1270
Vincolo sul profitto	0	0	0
Vincolo sulla leva minima	0	0	0
Vincolo sulla leva massima	0	0	0
Vincolo di cardinalità	0	0	0
Vincolo sulla non negatività	0	0	0
Vincolo d'integralità	0	0	0
Vincolo sulla quota minima	0	0	0
Vincolo sulla quota massima	0	0	0

Tabella 11: risultati ottenuti da tre prove della PSO.

[53] Il periodo *out-of-sample* è dal 02/01/2014 al 02/04/2014.

Prova	***1***	***2***	***3***
Fitness	2560.6	2227.2	3802.3
Misura di rischio out-of-sample	1257.6	1808.6	1834.8
Misura di rischio in-sample	2550	2227.2	3291.1
Profitti out-of-sample	2005.4	1269.2	2116.1
Profitti in-sample	1669.9	1295.2	1497.8
Vincolo sul profitto	0	0	0
Vincolo sulla leva minima	0	0	0
Vincolo sulla leva massima	0	0	0
Vincolo di cardinalità	0	0	0
Vincolo sulla non negatività	0	0	0
Vincolo d'integralità	0	0	0
Vincolo sulla quota minima	0	0	0
Vincolo sulla quota massima	0.0011	0	0.0451

Tabella 12: risultati ottenuti da tre prove della BFO.

Da una prima analisi dei dati si può notare come la BFO presenti una maggiore difficoltà a rispettare tutti i vincoli rispetto alla PSO. Inoltre, da un punto di vista sia *in-sample* che *out-of-sample*, la prima determina delle allocazioni che presentano una maggiore rischiosità rispetto alla seconda. Per quanto riguarda i profitti, dal lato *in-sample* la BFO è sempre superiore alla PSO; tale regola, ad eccezione di un singolo caso, vale anche per i profitti *out-of-sample*. Riassumendo, si potrebbe affermare che il maggior rischio connesso con le soluzioni ottenute attraverso la BFO, è giustificato da un valore più elevato dei profitti medi giornalieri.

Questo processo di calcolo dell'allocazione ottima ed utilizzo della stessa per il futuro virtuale è stato ripetuto altre diciassette volte fino al 29/06/2018. Come evidenziato nei capitoli precedenti, è impossibile sapere se l'allocazione ottenuta mediante le metaeuristiche sia un ottimo globale oppure solo vicina ad un ottimo globale: non è possibile conoscere neppure la distanza tra queste due soluzioni. Quindi, con il fine di mitigare questo limite, per determinare le varie allocazioni da utilizzare in tutti i periodi *out-of-sample* si è deciso di effettuare tre prove per ogni

settaggio considerato, per poi selezionare la soluzione che presenta la fitness con il valore più basso. Nelle seguenti tabelle sono riportati, per ogni anno e per ogni metaeuristica, i risultati ottenuti nei quattro periodi *out-of-sample* (ognuno costituito da 65 osservazioni). In particolare, tutte le righe che compongono ogni tabella mostrano i seguenti valori: i profitti medi giornalieri e la misura di rischio two-sided, entrambi realizzati in quello specifico intervallo temporale *out-of-sample*; il numero di vincoli violati dall'allocazione ottima utilizzata in quel determinato periodo; il profitto cumulato risultante dalla somma di tutti i profitti giornalieri ottenuti nel periodo considerato.

Periodo out-of-sample	***Profitto medio giornaliero***	***Misura di rischio***	***N° vincoli violati***	***Profitto cumulato***
02/01-02/04	1342.00	926.6	0	87227.5
03/04-02/07	-369.17	3504.3	0	-23996
03/07-01/10	759.81	1305.2	0	49387.75
02/10-31/12	473.40	2394.1	0	30771.25

Tabella 13: PSO-risultati out-of-sample nel 2014.

Periodo out-of-sample	***Profitto medio giornaliero***	***Misura di rischio***	***N° vincoli violati***	***Profitto cumulato***
02/01-02/04	1269.20	1808.6	0	82499.50
03/04-02/07	-772.76	4234.8	1	-50299.50
03/07-01/10	474.35	1882.6	0	30833.00
02/10-31/12	671.74	2345.6	0	43663.25

Tabella 14: BFO-risultati out-of-sample nel 2014.

Periodo out-of-sample	***Profitto medio giornaliero***	***Misura di rischio***	***N° vincoli violati***	***Profitto cumulato***
01/01-01/04	1379.40	2715.4	1	89367
02/04-01/07	-608.12	4521.5	1	-39528
02/07-30/09	-1200.00	5467	0	-78004.25
01/10-31/12	-72.27	1605.1	0	-4697.75

Tabella 15: PSO-risultati out-of-sample nel 2015.

Periodo out-of-sample	*Profitto medio giornaliero*	*Misura di rischio*	*N° vincoli violati*	*Profitto cumulato*
01/01-01/04	1048.20	2833.4	1	68130.50
02/04-01/07	-1731.50	7198.8	0	-112546.00
02/07-30/09	-703.73	4902.9	0	-45743.00
01/10-31/12	-66.48	2130.8	0	-4321.50

Tabella 16: BFO-risultati out-of-sample nel 2015.

Periodo out-of-sample	*Profitto medio giornaliero*	*Misura di rischio*	*N° vincoli violati*	*Profitto cumulato*
01/01-31/03	-47.93	3825.9	2	-3116
01/04-30/06	1646.10	1777.8	0	106996.25
01/07-29/09	-433.47	5254.2	0	-28175.75
30/09-30/12	-280.36	4733.2	0	-18223.5

Tabella 17: PSO-risultati out-of-sample nel 2016.

Periodo out-of-sample	*Profitto medio giornaliero*	*Misura di rischio*	*N° vincoli violati*	*Profitto cumulato*
01/01-31/03	-612.99	4506.1	1	-39844.50
01/04-30/06	1662.60	1494.9	1	108066.50
01/07-29/09	528.62	3600.4	0	34360.50
30/09-30/12	-420.54	6172	0	-27335.25

Tabella 18: BFO-risultati out-of-sample nel 2016.

Periodo out-of-sample	*Profitto medio giornaliero*	*Misura di rischio*	*N° vincoli violati*	*Profitto cumulato*
02/01-31/03	628.34	1988.8	0	40842.5
03/04-30/06	222.81	1597.5	0	14482.75
03/07-29/09	-69.91	2375	1	-4544.75
02/10-29/12	-38.59	2274.7	0	-2508.25

Tabella 19: PSO-risultati out-of-sample nel 2017.

Periodo out-of-sample	*Profitto medio giornaliero*	*Misura di rischio*	*N° vincoli violati*	*Profitto cumulato*
02/01-31/03	652.20	2650.9	0	42393.00
03/04-30/06	657.93	1610.4	0	42765.50
03/07-29/09	273.35	1979.4	0	17767.75
02/10-29/12	201.73	3052.7	1	13112.50

Tabella 20: BFO-risultati out-of-sample nel 2017.

Periodo out-of-sample	*Profitto medio giornaliero*	*Misura di rischio*	*N° vincoli violati*	*Profitto cumulato*
01/01-30/03	233.07	3358.9	0	15149.5
02/04-29/06	13.06	2396.7	0	848.75

Tabella 21: PSO-risultati out-of-sample nel 2018.

Periodo out-of-sample	*Profitto medio giornaliero*	*Misura di rischio*	*N° vincoli violati*	*Profitto cumulato*
01/01-30/03	345.06	2834.4	0	22428.75
02/04-29/06	231.49	1956.5	0	15046.75

Tabella 22: BFO-risultati out-of-sample nel 2018.

A questo punto è possibile procedere con la scelta dell'algoritmo che verrà utilizzato nel seguito per la determinazione dell'allocazione finale. Osservando i risultati ottenuti nei differenti futuri virtuali, non si può affermare con certezza che una metaeuristica sia nettamente superiore rispetto all'altra. La BFO ha determinato allocazioni che hanno fornito un profitto giornaliero medio di 206.03 dollari contro i 198.79 dollari della PSO. Tuttavia, per quanto riguarda il rischio, la BFO porta ad investire in allocazioni più rischiose rispetto alla PSO: il valore medio della misura di rischio two-sided è stato 3177.51 per la prima e 2890.11 per la seconda. A riprova di ciò si osservino le Figure 26 e 27, nelle quali si può notare che la distribuzione dei profitti *out-of-sample*, realizzati attraverso l'utilizzo della BFO, presenta degli indici di curtosi ed asimmetria superiori rispetto agli stessi ottenuti mediante la PSO.

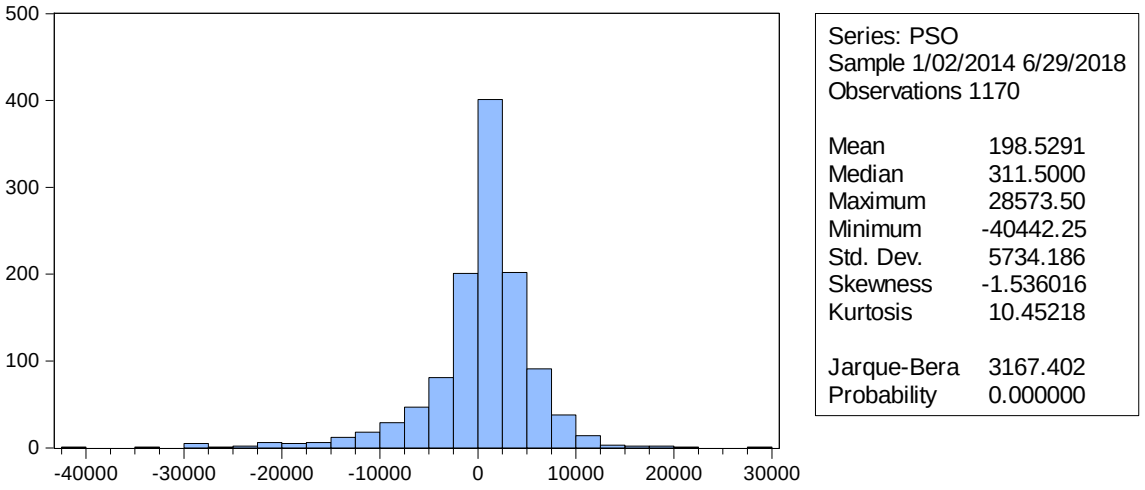


Figura 26: PSO-istogramma ed analisi descrittive dei profitti out-of-sample.

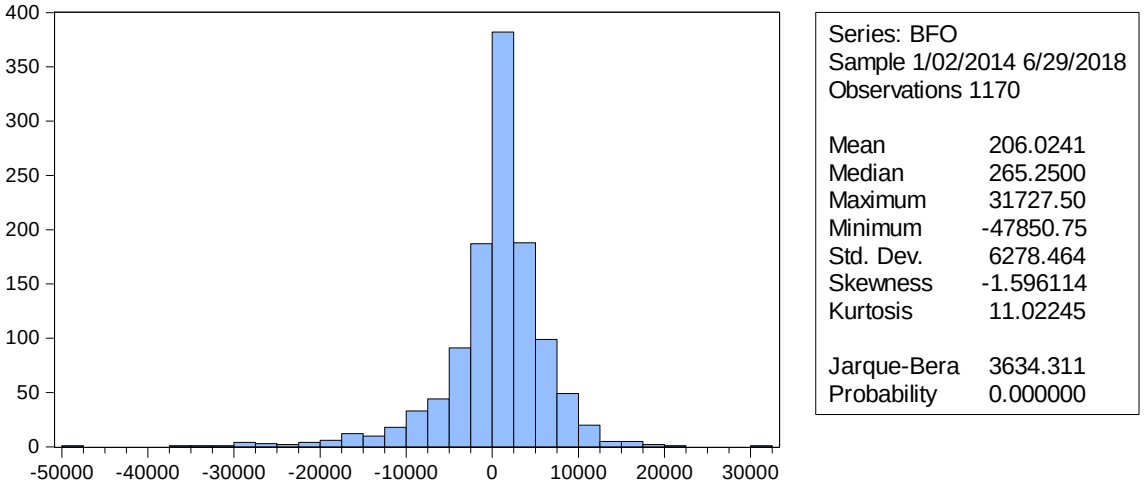


Figura 27: BFO-istogramma ed analisi descrittive dei profitti out-of-sample.

Per quanto riguarda la violazione dei vincoli, nel complesso le due metaeuristiche si sono comportate in modo molto simile.

Con il fine di determinare la metaeuristica migliore (nell'ipotesi che esista) tra la BFO e la PSO, si è fatto ricorso all'utilizzo di una misura di performance corretta per il rischio, ovvero l'indice di Sortino. Tale indice è espresso come:

$$Sortino = \frac{\hat{r}_p - S}{DSR},$$

dove:

- S è un rendimento soglia scelto dall'investitore;
- $DSR = \sqrt{Var\left[min\left(0, \hat{r}_p - S\right)\right]}$, ovvero il *Downside Risk*.

Questo può essere definito come un indice che misura il livello di performance generata a fronte di una determinata quantità di rischio di perdita assunta.

In queste analisi, il rendimento medio è sostituito con il profitto medio dell'allocazione ottima, mentre come soglia S è stato scelto lo zero. La Tabella 23 mostra i risultati conseguiti per entrambe le metaeuristiche.

Metaeuristica	Profitto medio	DSR	Indice di Sortino
PSO	198.79	4206.078	0.047263
BFO	206.03	4610.659	0.044686

Tabella 23: profitto medio, DSR ed indice di Sortino ottenuti dai dati out-of-sample.

Come si può vedere, i risultati ottenuti mediante la PSO presentano un indice di Sortino leggermente superiore rispetto allo stesso ricavato utilizzando la BFO.[54] Di conseguenza, si è deciso di utilizzare la PSO per la determinazione dell'allocazione finale nel programma Styx, da consigliare a DeltaHedge.

5.5 Determinazione dell'allocazione finale

Dopo aver determinato la metaeuristica che offre i risultati migliori a livello *out-of-sample*, non rimane che effettuare l'ultimo fondamentale passo di questo lavoro, ovvero individuare l'allocazione ottima per il programma Styx, la quale verrà poi consigliata a DeltaHedge. A tale scopo sono stati utilizzati, come dati *in-sample*, i profitti giornalieri dal 10/07/2017 al 10/07/2018. Attraverso la PSO sono state effettuate cinque prove, ottenendo cinque allocazioni differenti. La Tabella 24 mostra gli output ottenuti per ogni prova, mentre la Tabella 25 figura le diverse all'allocazioni associate ad ogni prova.

Prova	***1***	***2***	***3***	***4***	***5***
Fitness	2121.7	2087.1	2098.2	2127.1	2120.5
Misura di rischio in-sample	2121.7	2087.1	2098.2	2127.1	2120.5
Profitti in-sample	149.64	116.5239	141.9857	143.93	143.9055
Vincolo sul profitto	0	0	0	0	0
Vincolo sulla leva minima	0	0	0	0	0
Vincolo sulla leva massima	0	0	0	0	0
Vincolo di cardinalità	0	0	0	0	0

[54] È stata effettuata la stessa prova utilizzando una seconda misura di performance corretta per il rischio, ovvero l'indice di Sharpe, ottenendo dei risultati in linea con quelli dell'indice di Sortino (ossia 0.03468 per la PSO e 0.03283 per la BFO).

Vincolo sulla non negatività	0	0	0	0	0
Vincolo d'integralità	0	0	0	0	0
Vincolo sulla quota minima	0	0	0	0	0
Vincolo sulla quota massima	0	0	0	0	0

Tabella 24: output cinque prove PSO.

Prova	***6A***	***6B***	***6C***	***6E***	***6J***	***6M***	***6N***	***6S***	***DX***
1	2	3	2	1	3	2	7	2	1
2	1	3	1	0	4	2	7	2	3
3	1	3	2	1	3	1	7	2	2
4	1	3	1	1	4	0	7	2	2
5	1	3	2	1	2	2	7	2	3

Tabella 25: allocazioni ottenute per ogni prova.

È interessante vedere come le cinque allocazioni, ottenute attraverso le cinque prove, siano piuttosto simili: in tutti i casi sono stati attribuiti tre contratti nella strategia sul 6B, sette contratti nella strategia sul 6N e due contratti nella strategia sul 6S. Inoltre, sono tutte applicabili a livello pratico, in quanto non violano alcun vincolo. La scelta dell'allocazione ottima finale è ricaduta su quella relativa alla seconda prova (evidenziata in giallo), poiché presenta un valore della fitness inferiore rispetto a quello ottenuto negli altri quattro casi. Il valore della leva finanziaria, connesso con questa allocazione, è dello 0.955698, quindi si posiziona sul lato sinistro all'interno dell'intervallo di leva consentito. Quest'ultimo elemento può essere considerato positivamente in quanto, a parità di fitness, è preferibile utilizzare un'allocazione che sfrutti un livello di leva più basso.

Il numero elevato di contratti futures associati alla strategia sul 6N è giustificato dalle ottime performance realizzate nel periodo *in-sample* da tale strategia. Tuttavia, si rende necessario evidenziare che i futures sul tasso di cambio tra il dollaro Neozelandese e il dollaro Americano non sono molto liquidi, quindi, nel caso si aumentasse il livello di capitale investito, sarebbe preferibile introdurre un vincolo

sulla liquidità.

Un'ultima osservazione riguarda i due parametri della misura di rischio two-sided utilizzati. Questi sono stati mantenuti costanti per tutte le analisi effettuate, ipotizzando un'attitudine al rischio di DeltaHedge che non cambia nel tempo. Una possibile formulazione alternativa potrebbe essere quella di modificare i due parametri p ed α sulla base della variazione della volatilità dei mercati in un determinato periodo: individuare le allocazioni ottime attraverso un elevato valore di p ed un basso valore di α quando i mercati sono caratterizzati da un'elevata volatilità; calcolare le allocazioni ottime sulla base di un basso valore di p ed un elevato valore di α in periodi caratterizzati da una bassa volatilità dei mercati. In altre parole, tale approccio si fonda sull'ipotesi che l'attitudine al rischio del singolo investitore sia dinamica, in quanto si modifica all'aumentare o al diminuire della volatilità dei mercati.

Conclusioni

Nel presente lavoro è stata determinata un'allocazione "ottima", per il *managed futures program* Styx, da consigliare alla *Commodity Trading Association* che lo gestisce, ovvero DeltaHedge. In altri termini, sono stati determinati il numero di contratti futures da associare ad ognuna delle nove strategie di trading contrarian, sulle quali il programma investe 2 milioni di dollari. La Tabella 26 mostra l'allocazione finale che verrà fornita a DeltaHedge.

Strategia	**6A**	**6B**	**6C**	**6E**	**6J**	**6M**	**6N**	**6S**	**DX**
Numero contratti	1	3	1	0	4	2	7	2	3

Tabella 26: allocazione "ottima" finale di Styx.

Tale allocazione, che può essere definita come soluzione del problema di minimizzazione proposto da DeltaHedge, è stata determinata utilizzando una metaeuristica, ossia la PSO. La scelta di utilizzare quest'ultima è il risultato derivante da una preventiva analisi, dove questa è stata confrontata con la BFO (un'altra metaeuristica *bio-inspired*). Le analisi svolte mostrano come la BFO conduca ad allocazioni che presentano un valore *out-of-sample* della misura di rischio più elevato rispetto allo stesso ottenuto nel caso della PSO. Tuttavia, tale maggiore rischiosità è compensata da un valore più elevato dei profitti medi giornalieri *out-of-sample*. Per quanto riguarda i vincoli del modello, entrambe le metaeuristiche, nella maggior parte dei casi, hanno individuato soluzioni che rispettano tutti gli otto vincoli impostati. Dato che nessuna delle due metaeuristiche ha mostrato un'evidente superiorità rispetto all'altra, si è fatto ricorso a una misura di performance corretta per il rischio, ossia l'indice di Sortino, il quale è risultato leggermente superiore per la PSO.

Naturalmente, l'allocazione ottima finale in Tabella 26 non deve essere mantenuta per un tempo indefinito, ma deve essere continuamente revisionata dopo un intervallo di tempo predefinito (ad esempio ogni tre mesi).

Un'ultima osservazione riguarda la liquidità dei contratti. Infatti, nel momento in cui si andasse ad aumentare il capitale a disposizione del fondo, e quindi il numero di contratti associati ad ogni strategia, sarebbe preferibile aggiungere un vincolo al

modello di selezione con il fine di obbligare la PSO a fornire allocazioni che rispettano un predeterminato valore di uno specifico indice di liquidità. Quest'ultimo consentirebbe di ridurre lo *slippage* negativo, ovvero i costi opportunità scaturenti dall'eseguire un determinato ordine di acquisto (o vendita) ad un prezzo superiore (o inferiore) rispetto a quello a cui doveva essere effettivamente eseguito.

Appendice A

Legenda sigle futures:

6A: Australian Dollar Future
6B: British Pound Future
6C: Canadian Dollar Future
6E: Euro FX Future
6J: Japanese Yen Future
6M: Mexican Peso Future
6N: New Zealand Dollar Future
6S: Swiss Franc Future
DX: Dollar Index Future

Appendice B

Legenda sigle cambi:

AUD: Dollaro Australiano

CAD: Dollaro Canadese

CHF: Franco Svizzero

DX: Dollar Index

EUR: Euro

GBP: Sterlina Inglese

JPY: Yen

MXN: Peso Messicano

NZD: Dollaro Neozelandese

USD: Dollaro Americano

Appendice C

Codice Matlab: PSO.

```
clear all;
close all;
clc;

format short

%caricamento dati
load Scenario.csv; %importa i profitti
profitti = Scenario;
[n,numvar] = size(profitti);

before = 261; %numero di profitti in-smaple
before = before + 1;
oos = 65; %numero di profitti out-of-sample
pr1 = profitti(1:before,:);
pr2 = profitti((before+1):n,:);
dim1 = size(pr1);
dim2 = size(pr2);
prj = profitti(1:before,5); %profitti giornalieri 6j
prdx = profitti(1:before,9); %profitti giornalieri DX
prjm = mean(prj); %profitto medio giornaliero 6j
prdxm = mean(prdx); %profitto medio giornaliero DX

prm1 = mean(pr1);
prm2 = mean(pr2);
Vc = [74590 82975 76270 146800 112612 26415 68380
124000 93890];
Ct = 2000000;

%inizializzazione parametri PSO
P = 30; %numero particelle
niter = 3000; %numero iterazioni
Kd = 7; %numero minimo di titoli detenibili
perc_min = ones(1,numvar)*0.01; % quota minima
perc_max = ones(1,numvar)*0.31; % quota massima
c1 = 1.49618;
c2 = 1.49618;
w = 0.7298;
vmaxx = zeros(1,numvar);
vmaxz = zeros(1,numvar);
%inizializzazione altri parametri
```

```
epsilon1 = 0.0001;
a = 0.5;
b = 2;
prm1s = sum(prm1);
pr = (prm1s*4)-(prjm*4)-(prdxm*4);
Lmin = 0.95; %leva minima
Lmax = 1.25; %leva massima

TT = (n-oos);
%vettori di appoggio per la funzione obiettivo
rho = zeros(P,1); %misura di rischio
prm_port = zeros(P,1); %mean of portfolio
p_1 = zeros(P,1);
p_2 = zeros(P,1);
vinc_1 = zeros(P,1); %vincolo sul rendimento minimo
atteso
vinc_2 = zeros(P,1); %vincolo sul numero minimo di
titoli (K)
vinc_3 = zeros(P,1); %vincolo sulla leva minima
vinc_4 = zeros(P,1); %vincolo sulla leva massima
app_1 = zeros(P,numvar);
vinc_5 = zeros(P,1); %x >= perc_min
app_2 = zeros(P,numvar);
vinc_6 = zeros(P,1); %x <= perc_max
app_3 = zeros(P,numvar);
vinc_7 = zeros(P,1); %z è 0 o è 1
app_4 = zeros(P,numvar);
vinc_8 = zeros(P,1); %x non negativo


% 1) inizializzazione posizioni, velocità, funzione
obiettivo
x = rand(P,numvar)*10;
vx = rand(P,numvar)*10;
z = rand(P,numvar);
vz = rand(P,numvar);
f = ones(P,1)*1.0e+015;
x1 = zeros(P,numvar); %matrice di appoggio che tiene
conto dell'effettiva presenza del titolo in
portafoglio

pbx = [x f];
pbz = z;

% g=gbest, vettore che rappresenta la migliore
posizione globale e il valore della funzione obiettivo
associata
```

```
gx = zeros(1,numvar+1);
gz = zeros(1,numvar);

differenza_1 = zeros(TT,numvar);
prodotto_somma_1 = zeros(TT,P);
differenza_2 = zeros(oos,numvar);
prodotto_somma_2 = zeros(oos,P);

tic;
for k = 1:niter
    %individua range dinamico velocità
    for i = 1:numvar
        vmaxx(i) = abs(max(x(:,i))-min(x(:,i)));
        vmaxz(i) = abs(max(z(:,i))-min(z(:,i)));
    end;

    Cp = x*Vc';
    % 2)calcola funzione obiettivo
    for p = 1:P
        for i = 1:numvar
            x1(p,i) = x(p,i)*z(p,i);
            app_1(p,i) = max(0,(perc_min(i)*z(p,i))-
((Vc(i)*x(p,i)))/Cp(p));
            app_2(p,i) = max(0,((Vc(i)*x(p,i))/Cp(p))-
(perc_max(i)*z(p,i)));
            app_3(p,i) = abs(z(p,i)*(1-z(p,i)));
            app_4(p,i) = max(0,-x(p,i));
        end;

    for t = 1:TT
       differenza_1(t,:) = pr1(t,:)- prm1;
       prodotto_somma_1(t,p) =
differenza_1(t,:)*x1(p,:)';
    end;
    for t = 1:oos
       differenza_2(t,:) = pr2(t,:)- prm2;
       prodotto_somma_2(t,p) =
differenza_2(t,:)*x1(p,:)';
    end;

    prm_port(p) = x1(p,:)*prm1';

        rho(p) =
((a/TT)*(sum(max(0,prodotto_somma_1(:,p))))+(1-
a)*(((1/TT)*(sum((max(0,-
prodotto_somma_1(:,p))).^b)))^(1/b)) - prm_port(p));
%misura di rischio 2 sided
```

```
        vinc_1(p) = max(0,pr-(sum(x(p,:)*prm1')));
%vincolo di reddittività
        vinc_2(p) = max(0,Kd-sum(z(p,:))); %min num
titoli K
        vinc_3(p) = max(0,Lmin-((Vc*x(p,:)')/Ct));
%vicolo leva minima, forse X1(p,:)'
        vinc_4(p) = max(0,((Vc*x(p,:)')/Ct) - Lmax);
%vicolo leva massima
        vinc_5(p) = sum(app_1(p,:)); %quota min
        vinc_6(p) = sum(app_2(p,:)); %quota max
        vinc_7(p) = sum(app_3(p,:)); %z è 0 o 1
        vinc_8(p) = sum(app_4(p,:)); %x non negativo

    end;

    %funzione obiettivo
    f = rho + ((1/epsilon1)*(vinc_1 + vinc_2 + vinc_3
+ vinc_4 + vinc_5 + vinc_6 + vinc_7 + vinc_8));

    % 3) 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) = x(p,i);
                pbz(p,i) = z(p,i);
            end;
        end;
    end;

    % 4)identifica la particella con migliore
posizione
    [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;

    % 5)aggiorna velocità e posizione
    for p = 1:P
        for i = 1:numvar
            vx(p,i) = w*vx(p,i)+c1*rand*(pbx(p,i)-
x(p,i))+c2*rand*(gx(i)-x(p,i));
            vz(p,i) = w*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;
            z(p,i) = z(p,i)+vz(p,i);
            x(p,i) = x(p,i)+vx(p,i);
        end;
    end;
    x = round(x); %cosi ottengo numeri interi
    z = round(z); %cosi ottengo numeri interi
    converg(k,:) = gx(:,end);

    % 6)riparti dal 2) fino al criterio di stop
end;

toc;

figure;
plot(converg),title('FITNESS
FUNCTION'),xlabel('Iteration'),ylabel('Fitness');

contratti_ottimi = gx(1,1:numvar)'
best_fitness = gx(1,numvar+1)
rho_1 = rho(posizione)
rho_2 = (a/(n-
TT))*(sum(max(0,differenza_2*contratti_ottimi)))+(1-
a)*(((1/(n-TT))*(sum(max(0,-
(differenza_2*contratti_ottimi)).^b)))^(1/b)) -
(contratti_ottimi'*prm2')

profitti_out_of_sample = gx(1,1:numvar)*prm2'
profitti_in_sample = gx(1,1:numvar)*prm1'
num_titoli = sum(z(posizione,:))

vinc_profitto = vinc_1(posizione)
vinc_k = vinc_2(posizione)
vinc_levamin = vinc_3(posizione)
vinc_levamax = vinc_4(posizione)
vinc_quotamin = vinc_5(posizione)
vinc_quotamax = vinc_6(posizione)
vinc_z = vinc_7(posizione)
vinc_x = vinc_8(posizione)
```

Codice Matlab: BFO.

```
clear all;
close all;
clc;

format short

%caricamento dati
load Scenario.csv; %importa i profitti
profitti = Scenario;
[n,numvar] = size(profitti);

before = 261; %numero di profitti per le analisi in-
sample
before = before + 1;
oos = 65; %numero di profitti per le analisi out-of-
sample
pr1 = profitti(1:before,:);
pr2 = profitti((before+1):n,:);
dim1 = size(pr1);
dim2 = size(pr2);
prj = profitti(1:before,5); %profitti giornalieri 6j
prdx = profitti(1:before,9); %profitti giornalieri DX
prjm = mean(prj); %profitto medio gironaliero 6j
prdxm = mean(prdx); %profitto medio gironaliero DX

prm1 = mean(pr1);
prm2 = mean(pr2);
Vc = [76790 87612 77550 154037 95437 27530 72360
119262 89810];
Ct = 2000000;

%inizializzazione parametri BFO
P = 30; %numero batteri
Kd = 7; %numero minimo di titoli detenibili
perc_min = ones(1,numvar)*0.01; % quota minima
perc_max = ones(1,numvar)*0.31; % quota massima
%inizializzazione altri parametri

epsilon1 = 0.0001;
a = 0.5;
b = 2;
prm1s = sum(prm1);
pr = (prm1s*4)-(prjm*4)-(prdxm*4);
Lmin = 0.95; %leva minima
Lmax = 1.25; %leva massima
```

```
TT = (n-oos);
Nc = 400; %numero fasi di chemiotassi
Ns = 4; %numero fasi di avanzamento
Nre = 4; %numero fasi di riproduzione
Ned = 4; %numero fasi di elimination-dispersal
Ped = 0.3; %probabilità di eliminazione
Ctu = ones(1,numvar)*0.02; %ampiezza step di
capovolgimento
Cs = ones(1,numvar)*0.0002; %ampiezza step di
avanzamento
Cp = ones(P,1);

%vettori di appoggio per la funzione obiettivo
rho = zeros(P,1); %misura di rischio
prm_port = zeros(P,1); %mean of portfolio
p_1 = zeros(P,1);
p_2 = zeros(P,1);
vinc_1 = zeros(P,1); %vincolo sul rendimento minimo
atteso
vinc_2 = zeros(P,1); %vincolo sul numero minimo di
titoli (K)
vinc_3 = zeros(P,1); %vincolo sulla leva minima
vinc_4 = zeros(P,1); %vincolo sulla leva massima
app_1 = zeros(P,numvar);
vinc_5 = zeros(P,1); %x >= perc_min
app_2 = zeros(P,numvar);
vinc_6 = zeros(P,1); %x <= per_max
app_3 = zeros(P,numvar);
vinc_7 = zeros(P,1); %z è 0 o è 1
app_4 = zeros(P,numvar);
vinc_8 = zeros(P,1); %x non negativo

%inizializzazione posizioni, velocità, funzione
obiettivo
x = rand(P,numvar)*10;
xtu = rand(P,numvar)*10;
ztu = rand(P,numvar);
xed = rand(P,numvar)*10;
zed = rand(P,numvar);
z = rand(P,numvar);
f = ones(P,1)*1.0e+015;
ftu = ones(P,1)*1.0e+015;
fed = ones(P,1)*1.0e+015;
x1 = zeros(P,numvar); %matrice di appoggio che tiene
conto dell'effettiva presenza del titolo in
portafoglio
```

```
differenza_1 = zeros(TT,numvar);
prodotto_somma_1 = zeros(TT,P);
differenza_2 = zeros(oos,numvar);
prodotto_somma_2 = zeros(oos,P);


    %calcola funzione obiettivo
    for p = 1:P
       Cp(p) = x(p,:)*Vc';
        for i = 1:numvar
            x1(p,i) = x(p,i)*z(p,i);
            app_1(p,i) = max(0,(perc_min(i)*z(p,i))-
((Vc(i)*x(p,i)))/Cp(p));
            app_2(p,i) = max(0,((Vc(i)*x(p,i))/Cp(p))-
(perc_max(i)*z(p,i)));
            app_3(p,i) = abs(z(p,i)*(1-z(p,i)));
            app_4(p,i) = max(0,-x(p,i));
        end;

    for t = 1:TT
       differenza_1(t,:) = pr1(t,:)- prm1;
       prodotto_somma_1(t,p) =
differenza_1(t,:)*x1(p,:)';
    end;
    for t = 1:(n-TT)
       differenza_2(t,:) = pr2(t,:)- prm2;
       prodotto_somma_2(t,p) =
differenza_2(t,:)*x1(p,:)';
    end;

    prm_port(p) = x1(p,:)*prm1';

        rho(p) =
((a/TT)*(sum(max(0,prodotto_somma_1(:,p))))+(1-
a)*(((1/TT)*(sum((max(0,-
prodotto_somma_1(:,p))).^b)))^(1/b)) - prm_port(p));
%misura di rischio 2 sided
        vinc_1(p) = max(0,pr-(sum(x(p,:)*prm1')));
%vincolo di reddittività
        vinc_2(p) = max(0,Kd-sum(z(p,:))); %min num
titoli K
        vinc_3(p) = max(0,Lmin-((Vc*x(p,:)')/Ct));
%vicolo leva minima
        vinc_4(p) = max(0,((Vc*x(p,:)')/Ct) - Lmax);
%vicolo leva massima
        vinc_5(p) = sum(app_1(p,:)); %quota min
        vinc_6(p) = sum(app_2(p,:)); %quota max
```

```
        vinc_7(p) = sum(app_3(p,:)); %z è 0 o 1
        vinc_8(p) = sum(app_4(p,:)); %x non negativo

    end;

    %funzione obiettivo
    f = rho + ((1/epsilon1)*(vinc_1 + vinc_2 + vinc_3
+ vinc_4 + vinc_5 + vinc_6 + vinc_7 + vinc_8));

flast = f;
xlast = x;
zlast = z;
rholast = rho;

tic;

for ned = 1:Ned
  for nre = 1:Nre
    for nc = 1:Nc %step di chemiotassi

      D = randn(1,numvar); %numero casuale che
definisce l'ammontare del cambiamento di direzione
dovuto al capovolgimento
         for p = 1:P
           for i = 1:numvar
                xtu(p,i) = x(p,i) + Ctu(1,i)*D(1,i);
%calcolo il nuovo x dopo il capovolgimento
                ztu(p,i) = z(p,i) + Ctu(1,i)*D(1,i);
                if xtu(p,i) < 0
                    xtu(p,i) = 0;
                end
                if ztu(p,i) < 0
                    ztu(p,i) = 0;
                end
           end;
         end;
         %calcola funzione obiettivo
    for p = 1:P
        Cp(p) = xtu(p,:)*Vc';
        for i = 1:numvar
            x1(p,i) = xtu(p,i)*ztu(p,i);
            app_1(p,i) = max(0,(perc_min(i)*ztu(p,i))-
((Vc(i)*xtu(p,i)))/Cp(p));
            app_2(p,i) =
max(0,((Vc(i)*xtu(p,i))/Cp(p))-
(perc_max(i)*ztu(p,i)));
            app_3(p,i) = abs(ztu(p,i)*(1-ztu(p,i)));
            app_4(p,i) = max(0,-xtu(p,i));
```

```
        end;

    for t = 1:TT
       differenza_1(t,:) = pr1(t,:)- prm1;
       prodotto_somma_1(t,p) =
differenza_1(t,:)*x1(p,:)';
    end;
    for t = 1:(n-TT)
       differenza_2(t,:) = pr2(t,:)- prm2;
       prodotto_somma_2(t,p) =
differenza_2(t,:)*x1(p,:)';
    end;

    prm_port(p) = x1(p,:)*prm1';

        rho(p) =
((a/TT)*(sum(max(0,prodotto_somma_1(:,p))))+(1-
a)*(((1/TT)*(sum((max(0,-
prodotto_somma_1(:,p))).^b)))^(1/b)) - prm_port(p));
%misura di rischio 2 sided
        vinc_1(p) = max(0,pr-(sum(xtu(p,:)*prm1')));
%vincolo di reddittività
        vinc_2(p) = max(0,Kd-sum(ztu(p,:))); %min num
titoli K
        vinc_3(p) = max(0,Lmin-((Vc*xtu(p,:)')/Ct));
%vicolo leva minima
        vinc_4(p) = max(0,((Vc*xtu(p,:)')/Ct) - Lmax);
%vicolo leva massima
        vinc_5(p) = sum(app_1(p,:)); %quota min
        vinc_6(p) = sum(app_2(p,:)); %quota max
        vinc_7(p) = sum(app_3(p,:)); %z è 0 o 1
        vinc_8(p) = sum(app_4(p,:)); %x non negativo

    end;

    %funzione obiettivo
    ftu = rho + ((1/epsilon1)*(vinc_1 + vinc_2 +
vinc_3 + vinc_4 + vinc_5 + vinc_6 + vinc_7 + vinc_8));


for p = 1:P
  if ftu(p) < flast(p) %confronto la nuova funzione
obiettivo con la precedente
     flast(p) = ftu(p);
     for i = 1:numvar
     x(p,i) = xtu(p,i);
     z(p,i) = ztu(p,i);
     xlast(p,i) = x(p,i);
```

```
      zlast(p,i) = z(p,i);
      end
      for ns = 1:Ns %inizio step di avanzamento
         for i = 1:numvar
            x(p,i) = x(p,i) + Cs(1,i)*D(1,i); %calcolo
il nuovo x dopo l'avanzamento
            z(p,i) = z(p,i) + Cs(1,i)*D(1,i);
            if x(p,i) < 0
                x(p,i) = 0;
             end;
            if z(p,i) < 0
                z(p,i) = 0;
            end;
        end;
Cp(p) = x(p,:)*Vc';
         for i = 1:numvar
             x1(p,i) = x(p,i)*z(p,i);
             app_1(p,i) = max(0,(perc_min(i)*z(p,i))-
((Vc(i)*x(p,i)))/Cp(p));
             app_2(p,i) = max(0,((Vc(i)*x(p,i))/Cp(p))-
(perc_max(i)*z(p,i)));
             app_3(p,i) = abs(z(p,i)*(1-z(p,i)));
             app_4(p,i) = max(0,-x(p,i));
         end;

    for t = 1:TT
       differenza_1(t,:) = pr1(t,:)- prm1;
       prodotto_somma_1(t,p) =
differenza_1(t,:)*x1(p,:)';
    end;
    for t = 1:(n-TT)
       differenza_2(t,:) = pr2(t,:)- prm2;
       prodotto_somma_2(t,p) =
differenza_2(t,:)*x1(p,:)';
    end;

    prm_port(p) = x1(p,:)*prm1';

         rho(p) =
((a/TT)*(sum(max(0,prodotto_somma_1(:,p))))+(1-
a)*(((1/TT)*(sum((max(0,-
prodotto_somma_1(:,p))).^b)))^(1/b)) - prm_port(p));
%misura di rischio 2 sided
         vinc_1(p) = max(0,pr-(sum(x(p,:)*prm1')));
%vincolo di reddittività
         vinc_2(p) = max(0,Kd-sum(z(p,:))); %min num
titoli K
```

```
        vinc_3(p) = max(0,Lmin-((Vc*x(p,:)')/Ct));
%vicolo leva minima
        vinc_4(p) = max(0,((Vc*x(p,:)')/Ct) - Lmax);
%vicolo leva massima
        vinc_5(p) = sum(app_1(p,:)); %quota min
        vinc_6(p) = sum(app_2(p,:)); %quota max
        vinc_7(p) = sum(app_3(p,:)); %z è 0 o 1
        vinc_8(p) = sum(app_4(p,:)); %x non negativo

    %funzione obiettivo
    f(p) = rho(p) + ((1/epsilon1)*(vinc_1(p) +
vinc_2(p) + vinc_3(p) + vinc_4(p) + vinc_5(p) +
vinc_6(p) + vinc_7(p) + vinc_8(p)));

                if f(p) < flast(p)
                    flast(p) = f(p);
                    for i = 1:numvar
                   xlast(p,i) = x(p,i);
                   zlast(p,i) = z(p,i);
                    end
                end;
          end;
       end;
    end;
 end; %fine chemiotassi

xf = [xlast flast]; %inizio step riproduzione
    xhealth = sortrows(xf,10);
    m = P/2;
    xm = xhealth(1:m,1:numvar);
    xlast = [xm; xm];

zf = [zlast flast];
     zhealth = sortrows(zf,10);
     zm = zhealth(1:m,1:numvar);
     zlast = [zm; zm];
  xlast = round(xlast); %per ottenere contratti interi
  zlast = round(zlast); %per ottenere contratti interi
  x = xlast;
  z = zlast;

for p = 1:P
        Cp(p) = x(p,:)*Vc';
        for i = 1:numvar
            x1(p,i) = x(p,i)*z(p,i);
            app_1(p,i) = max(0,(perc_min(i)*z(p,i))-
((Vc(i)*x(p,i)))/Cp(p));
```

```
            app_2(p,i) = max(0,((Vc(i)*x(p,i))/Cp(p))-
(perc_max(i)*z(p,i)));
            app_3(p,i) = abs(z(p,i)*(1-z(p,i)));
            app_4(p,i) = max(0,-x(p,i));
        end;

    for t = 1:TT
       differenza_1(t,:) = pr1(t,:)- prm1;
       prodotto_somma_1(t,p) =
differenza_1(t,:)*x1(p,:)';
    end;
    for t = 1:(n-TT)
       differenza_2(t,:) = pr2(t,:)- prm2;
       prodotto_somma_2(t,p) =
differenza_2(t,:)*x1(p,:)';
    end;

    prm_port(p) = x1(p,:)*prm1';

        rho(p) =
((a/TT)*(sum(max(0,prodotto_somma_1(:,p))))+(1-
a)*(((1/TT)*(sum((max(0,-
prodotto_somma_1(:,p))).^b)))^(1/b)) - prm_port(p));
%misura di rischio 2 sided
        vinc_1(p) = max(0,pr-(sum(x(p,:)*prm1')));
%vincolo di reddittività
        vinc_2(p) = max(0,Kd-sum(z(p,:))); %min num
titoli K
        vinc_3(p) = max(0,Lmin-((Vc*x(p,:)')/Ct));
%vicolo leva minima
        vinc_4(p) = max(0,((Vc*x(p,:)')/Ct) - Lmax);
%vicolo leva massima
        vinc_5(p) = sum(app_1(p,:)); %quota min
        vinc_6(p) = sum(app_2(p,:)); %quota max
        vinc_7(p) = sum(app_3(p,:)); %z è 0 o 1
        vinc_8(p) = sum(app_4(p,:)); %x non negativo

    end;

    %funzione obiettivo
    f = rho + ((1/epsilon1)*(vinc_1 + vinc_2 + vinc_3
+ vinc_4 + vinc_5 + vinc_6 + vinc_7 + vinc_8));


flast = f;
rholast = rho;
 vinc_1last = vinc_1;
 vinc_2last = vinc_2;
```

```
 vinc_3last = vinc_3;
 vinc_4last = vinc_4;
 vinc_5last = vinc_5;
 vinc_6last = vinc_6;
 vinc_7last = vinc_7;
 vinc_8last = vinc_8;

end; %fine step riproduzione

for p = 1:P %inizio step elimination-dispersal
     for i = 1:numvar
        r = rand;
        xed(p,i) = x(p,i);
        zed(p,i) = z(p,i);
        if r < Ped
            r = randn;
        xed(p,i) = r;
        zed(p,i) = r;
        if xed(p,i) < 0
                    xed(p,i) = 0;
                end
                if zed(p,i) < 0
                    zed(p,i) = 0;
                end
        end;
     end;
end;
 xed = round(xed); %per ottenere contratti interi
zed = round(zed); %per ottenere contratti interi

for p = 1:P
        Cp(p) = xed(p,:)*Vc';
        for i = 1:numvar
            x1(p,i) = xed(p,i)*zed(p,i);
            app_1(p,i) = max(0,(perc_min(i)*zed(p,i))-
((Vc(i)*xed(p,i)))/Cp(p));
            app_2(p,i) =
max(0,((Vc(i)*xed(p,i))/Cp(p))-
(perc_max(i)*zed(p,i)));
            app_3(p,i) = abs(zed(p,i)*(1-zed(p,i)));
            app_4(p,i) = max(0,-xed(p,i));
        end;

    for t = 1:TT
       differenza_1(t,:) = pr1(t,:)- prm1;
       prodotto_somma_1(t,p) =
differenza_1(t,:)*x1(p,:)';
```

```
    end;
    for t = 1:(n-TT)
       differenza_2(t,:) = pr2(t,:)- prm2;
       prodotto_somma_2(t,p) =
differenza_2(t,:)*x1(p,:)';
    end;

    prm_port(p) = x1(p,:)*prm1';

        rho(p) =
((a/TT)*(sum(max(0,prodotto_somma_1(:,p))))+(1-
a)*(((1/TT)*(sum((max(0,-
prodotto_somma_1(:,p))).^b)))^(1/b)) - prm_port(p));
%misura di rischio 2 sided
        vinc_1(p) = max(0,pr-(sum(xed(p,:)*prm1')));
%vincolo di reddittività
        vinc_2(p) = max(0,Kd-sum(zed(p,:))); %min num
titoli K
        vinc_3(p) = max(0,Lmin-((Vc*xed(p,:)')/Ct));
%vicolo leva minima
        vinc_4(p) = max(0,((Vc*xed(p,:)')/Ct) - Lmax);
%vicolo leva massima
        vinc_5(p) = sum(app_1(p,:)); %quota min
        vinc_6(p) = sum(app_2(p,:)); %quota max
        vinc_7(p) = sum(app_3(p,:)); %z è 0 o 1
        vinc_8(p) = sum(app_4(p,:)); %x non negativo

    end;

    %funzione obiettivo
    fed = rho + ((1/epsilon1)*(vinc_1 + vinc_2 +
vinc_3 + vinc_4 + vinc_5 + vinc_6 + vinc_7 + vinc_8));

for p = 1:P
          if fed(p) < flast(p)
              flast(p) = fed(p);
              rholast(p) = rho(p);
               vinc_1last(p) = vinc_1(p);
               vinc_2last(p) = vinc_2(p);
               vinc_3last(p) = vinc_3(p);
               vinc_4last(p) = vinc_4(p);
               vinc_5last(p) = vinc_5(p);
               vinc_6last(p) = vinc_6(p);
               vinc_7last(p) = vinc_7(p);
               vinc_8last(p) = vinc_8(p);
              for i = 1:numvar
              x(p,i) = xed(p,i);
```

```
                z(p,i) = zed(p,i);
               xlast(p,i) = x(p,i);
               zlast(p,i) = z(p,i);
                end
            end;
end;

end; %fine step elimination-dispersal

   xf = [xlast flast rholast vinc_1last vinc_2last
vinc_3last vinc_4last vinc_5 vinc_6last vinc_7last
vinc_8last];
    xhealth = sortrows(xf,10);
    contratti = xhealth(1,1:9)'
    profitto_giornaliero_medio_is = prm1*contratti
    profitto_giornaliero_medio_oos = prm2*contratti
    rho_is = xhealth(1,11)
    rho_oos = ((a/(n-
TT))*(sum(max(0,differenza_2*contratti))))+(1-
a)*(((1/(n-TT))*(sum(max(0,-
(differenza_2*contratti)).^b)))^(1/b)) -
(contratti'*prm2')
    fitness = xhealth(1,10)
    zf = [zlast flast];
    zhealth = sortrows(zf,10);
    vincolo_z = zhealth(1,1:9)
    vinc_profitto = xhealth(1,12)
    vinc_k = xhealth(1,13)
    vinc_levamin = xhealth(1,14)
    vinc_levamax = xhealth(1,15)
    vinc_quotamin = xhealth(1,16)
    vinc_quotamax = xhealth(1,17)
    vinc_z = xhealth(1,18)
    vinc_x = xhealth(1,19)
```

Appendice D

Valori dei tassi di cambio utilizzati con il fine di determinare il valore totale di ogni singolo contratto futures alla fine di ogni periodo *in-sample*.

Data	6A	6B	6C	6E	6J	6M	6N	6S	DX
01/01/2014	0.8884	1.6566	0.9394	1.3754	0.9501	0.0766	0.8197	0.892	80.29
02/04/2014	0.9248	1.6627	0.9065	1.3767	0.9626	0.0763	0.8574	0.8867	80.37
02/07/2014	0.9444	1.7166	0.9376	1.366	0.9824	0.0769	0.8774	0.889	79.98
01/10/2014	0.8738	1.6186	0.8957	1.2624	0.9183	0.0743	0.7786	0.9559	86.08
31/12/2014	0.8170	1.5575	0.8605	1.2098	0.8354	0.0678	0.78	0.9942	90.65
01/04/2015	0.76	1.4823	0.7923	1.0764	0.8349	0.066	0.7452	0.9667	98.47
01/07/2015	0.7644	1.5618	0.7943	1.1054	0.8118	0.0633	0.6735	0.9483	96.51
30/09/2015	0.70	1.5129	0.751	1.1177	0.8343	0.0591	0.6395	0.9732	96.48
31/12/2015	0.7280	1.4739	0.7224	1.0861	0.8311	0.0582	0.6829	1.0021	98.35
31/03/2016	0.77	1.4362	0.7689	1.138	0.7805	0.0578	0.6909	0.9617	94.58
30/06/2016	0.75	1.3311	0.7736	1.1105	0.8717	0.0547	0.7134	0.9761	96.2
29/09/2016	0.76	1.297	0.7608	1.1223	0.8819	0.0512	0.7254	0.9661	95.47
30/12/2016	0.72	1.2338	0.7444	1.0516	0.8134	0.0482	0.6947	1.0184	102.29
31/03/2017	0.76	1.255	0.7511	1.0652	0.8426	0.0534	0.7007	1.0031	100.22
30/06/2017	0.77	1.3027	0.7713	1.1426	0.7787	0.0551	0.7333	0.9583	95.42
29/09/2017	0.78	1.3397	0.8019	1.1814	0.7524	0.0547	0.7207	0.9682	92.88
29/12/2017	0.78	1.3515	0.7949	1.1998	0.7396	0.0508	0.7087	0.9744	91.83
30/03/2018	0.77	1.4018	0.7755	1.2323	0.7635	0.0550	0.7236	0.9541	89.81
10/07/2018	0.75	1.3276	0.7627	1.1744	0.9009	0.0528	0.6838	0.992	93.89

Bibliografia

Acerbi, C., Tasche, D., 2002a. *Expected shortfall: a natural coherent alternative to value-at-risk*, Economic Notes by Banca Monte dei Paschi di Siena SpA 31 (2), 379-388.

Acerbi, C., Tasche, D., 2002b. *On the coherence of expected shortfall*, Journal of Banking and Finance 26 (7), 1487-1503.

Al Janabi, M., 2012. *Optimal commodity asset allocation with a coherent market risk modeling*, Review of Financial Economics 21, 131–140.

Artzner, P., Delbaen, F., Eber, J.M., Heath, D., 1999. *Coherent measures of risk*, Mathematical Finance 9, 203-228.

Basak, S., Shapiro, A., 2001. *Value-at-Risk-Based Risk Management: Optimal Policies and Asset Prices*, The Review of Financial Studies 14 (2), 371-405.

Blum, C., Roli, A., 2003. *Metaheuristics in Combinatorial Optimization: Overview and Conceptual Comparison*, ACM Computing Survey.

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

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

Cont, R., 2001. *Empirical properties of asset returns: stylized facts and statistical issues*, Quantitative finance 1, 223–236.

Corazza, M., Favaretto, D., 2007. *On the existence of solutions to the quadratic mixed-integer mean–variance portfolio selection problem*, European Journal of Operational Research 176 (3), 1947-1960.

Corazza, M., Fasano, G., Gusso, R., 2013. *Particle Swarm Optimization with non-smooth penalty reformulation, for a complex portfolio selection problem*, Applied Mathematics and Computation 224, 611-624.

Das, S., Biswas, A., Dasgupta, S., Abraham, A., 2009. *Bacterial Foraging Optimization Algorithm: Theoretical Foundations, Analysis, and Applications*, Foundations of Comput. Intel. 3, 23–55.

Kao, Y., Cheng, H. T., 2013. *Bacterial Foraging Optimization Approach to Portfolio Optimization*, Comput. Econ. 42, 453-470.

Laskari, E. C., Parsopoulos, K. E., Vrahatis, M. N., 2002. *Particle Swarm Optimization for Integer Programming*, Department of Mathematics, University of Patras Artificial Intelligence Research Center (UPAIRC), GR-26110 Patras, Greece.

Luenberger, D.G., 1984. *Linear and non linear programming*. Second edition, Addison-Wesley.

Mandelbrot, B., 1963. *New methods in statistical economics*, Journal of Political Economy 71 (5), 421-440.

Markowitz, H.M., 1952. *Portfolio selection*, Journal of Finance 7 (1), 77-91.

Mitton, T., Vorkink, K., 2007. *Equilibrium underdiversification and the preference for skewness*, The Review of Financial Studies 20 (4), 1255-1288.

Nadarajah, S., Zhang, B., Chan, S., 2014. *Estimation methods for expected shortfall*, Quantitative Finance 14 (2), 271–291.

Ortobelli, S., Rachev, S.T., Stoyanov, S., Fabozzi, F.J., Biglova, A., 2005. *The proper use of risk measures in portfolio theory,* International Journal of Theoretical and Applied Finance 8 (8), 1107–1133.

Peirò, A., 1999. *Skewness in financial returns*, Journal of Banking and Finance 23 (6), 847-862.

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

Rachev, S., Ortobelli, S., Stoyanov, S., Fabozzi, F. J., Biglova, A., 2008*. Desirable properties of an ideal risk measure in portfolio theory*, Int. J. Theor. Appl. Finan 11, 19-54.

Radovan, A., 2014. *La PSO per la selezione di portafogli basati sul Partitioned VaR e su vincoli misti-interi,* Università Ca' Foscari, Venezia.

Scott, R.C., Horvath, P.A., 1980. *On the direction of preference for moments of higher order than the variance*, Journal of Finance 35 (4), 915-919.

Szegö, G., 2002. *Measures of risk*, Journal of Banking and Finance 26 (7), 1253–1272.

Tchomté, S. K., Gourgand, M., 2009. *Particle swarm optimization: A study of particle displacement for solving continuous and combinatorial optimization problems*, Int. J. Production Economics 121, 57-67.

Trelea, I. C., 2003*. The particle swarm optimization algorithm: convergence analysis and parameter selection*, Information Processing Letters 85, 317–325.

Zarranezhad, M., Hajiabad, R. Y., Fakherinia, M., 2015. *Evaluation of particle swarm algorithm and genetic algorithms performance at optimal portfolio selection*, Asian Economic and Financial Review 5 (1), 88-101.