



Università Politecnica delle Marche

Dottorato in Economia Politica

XVI ciclo n.s.

**Asset Allocation e Copulae Multivariate
Dinamiche**

Andrea Faragalli

Supervisor:

Prof. Giulio Palomba

Coordinatore del corso:

Prof. Riccardo Lucchetti

Indice

1	Introduzione	5
2	Letteratura	15
2.1	Copulae	16
2.1.1	Famiglie di Copulae	24
2.1.2	Stima delle Copulae	29
2.1.3	Pair Copula Construction	31
2.1.4	R-vine copula	33
2.1.5	Stima modello R-vine a più stadi	39
2.1.6	Vantaggi e Svantaggi	43
2.2	Modelli di portafoglio con Copula	45
2.2.1	Patton (2004)	45
2.2.2	Riccetti (2013)	47
2.2.3	Truncated and simplified regular vines (Brechmann, Czado, and Aas, 2012)	49
2.2.4	Dynamic D-vine model (Almeida, Czado, and Manner, 2016)	51
2.3	Copulae e modelli GARCH	55
2.3.1	Modelli GARCH univariati	55
2.3.2	GARCH Multivariati	61
3	Modello di stima della struttura di dipendenza con copula multivariata dinamica	65

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

3.1	Modello R-vine dinamico	65
3.2	Stima del modello R-vine dinamico	69
3.3	Valutazione delle performance	75
3.3.1	Stime e performance in-sample	77
3.3.2	Previsioni	82
4	Ottimizzazione di portafoglio	87
4.0.1	Media-Varianza (Markowitz, 1952)	88
4.0.2	Global Minimum Variance	89
4.0.3	Media-Value at Risk (VaR)	91
4.0.4	Min-Conditional Value at Risk	92
4.1	Indici di performance di portafoglio	93
4.2	Portafoglio con copula R-vine dinamica	95
4.2.1	Risultati	97
5	Conclusioni	101
	Appendice	105
	Codice gretl	105
	Lista titoli utilizzati	147
	Statistiche descrittive	148

Capitolo 1

Introduzione

La continua evoluzione dei mercati finanziari ha portato, nel corso degli ultimi decenni, alla creazione di diverse tecniche per lo studio e la gestione dei rischi assunti dagli operatori finanziari nel corso delle loro attività di investimento. Il rischio di mercato, in ambito finanziario, è spesso inteso come la variabilità dei rendimenti degli strumenti finanziari nel corso del tempo.

L'elemento cardine di molti lavori empirici, è l'assunzione di normalità delle distribuzioni dei rendimenti dei diversi asset finanziari e della loro distribuzione congiunta (Markowitz (1952), Black and Scholes (1973)). Tale assunzione risulta, però, essere estremamente semplificatrice della realtà, infatti le serie dei rendimenti presentano spesso distribuzioni asimmetriche e leptocurtiche. Per tale motivo, la costruzione di portafogli diversificati, derivati da tecniche basate sulle ipotesi di normalità dei rendimenti, tendono in genere ad assumere dei rischi più elevati di quelli che tali modelli riescono a catturare. L'assunzione di normalità dei rendimenti comporta che vengano quasi ignorati gli eventi estremi che influenzano l'andamento degli indici finanziari, facendo sì che gli investitori si assumano un rischio più elevato di quello che realmente credono. Anche se ampiamente criticata e riconosciuta come estremamente semplificativa (Engle (1982)), l'assunzione di normalità ha trovato ampia applicazione empirica all'interno di tecniche di asset allocation e risk hedging. Ciò è dovuto

sia alla sua facile implementazione, infatti necessita di soli due momenti della distribuzione (Von Neumann and Morgenstern, 1947) di ogni singola serie di rendimenti e di un'unica assunzione riguardante la covarianza tra ogni coppia di asset, sia al fatto che essi ammettono delle soluzioni analitiche (si veda tra gli altri Palomba and Riccetti (2012)). I primi lavori di asset allocation si sono sviluppati prendendo in considerazione solo le esigenze dell'investitore (Markowitz, 1952). Nella realtà dei fatti gli investimenti vengono spesso gestiti da un asset manager o da un investitore istituzionale all'interno di fondi comuni di investimento. Richard Roll (1992) fu il primo a porre il problema della ricerca di portafogli efficienti che prendano in considerazione anche le esigenze dei manager responsabili degli investimenti attraverso l'imposizione di vincoli nella *Tracking Error Volatility*¹ (TEV). L'obiettivo di tale lavoro è la creazione di una *frontiera Media-TEV* all'interno del quale l'ottimizzazione di portafoglio viene effettuata attraverso la massimizzazione del rendimento medio rispetto al benchmark dato un vincolo sulla volatilità della TEV. Roll (1992) dimostra che tale ottimizzazione porta alla creazione di portafogli che sono sistematicamente più rischiosi rispetto al portafoglio di benchmark, e che essi non risultano neanche all'interno della frontiera efficiente di Markowitz (1952). In altre parole, il manager sceglierà un portafoglio che risulta essere dominato da altri portafogli, in quanto questi ultimi possiedono un rendimento maggiore ed una volatilità minore (sebbene non una minore tracking error volatility). Jorion (2003) estende questo modello e mostra che i portafogli con vincoli sulla TEV sono rappresentati da un'ellisse² all'interno del tradizionale spazio media-varianza. Viene messo in luce che aggiungendo un vincolo

¹Dato il rendimento di un portafoglio di riferimento (benchmark), essa misura la volatilità dello scostamento tra il rendimento del proprio portafoglio ed il rendimento del benchmark.

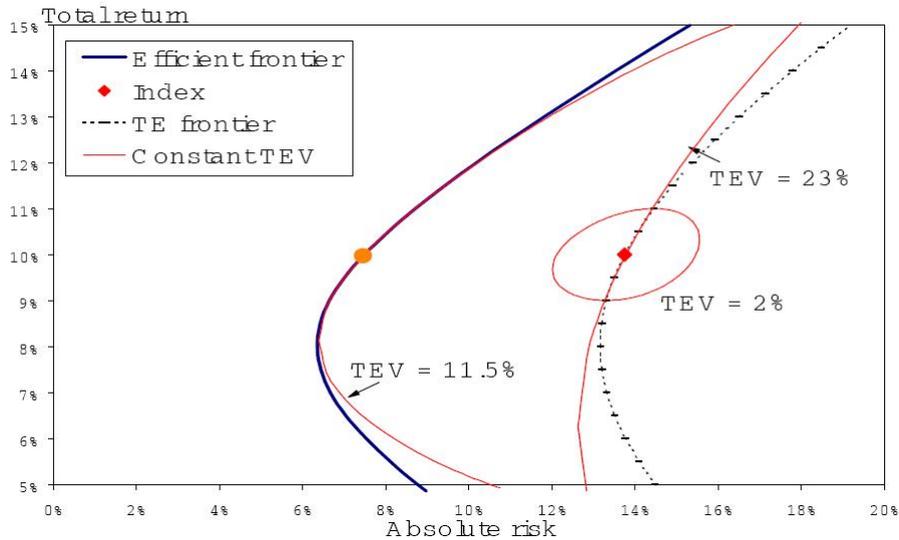
$$TEV = (w - w_b)'V(w - w_b),$$

dove $w - w_b$ rappresenta il vettore di deviazioni dal portafoglio benchmark e V la matrice di covarianza.

²Frontiera chiusa e limitata che si espande se la TEV aumenta, centrata sul Benchmark e con eccentricità che dipende dalla differenza tra μ_B e μ_p che sono rispettivamente il rendimento del Benchmark e il rendimento di portafogli

alla volatilità totale del portafoglio, si possono sostanzialmente migliorare le performance del portafoglio., Ulteriori studi hanno riguardo l'imposizione di

Figura 1.0.1: *Frontiera con vincoli sulla TEV*



NOTA: per Index si intende il portafoglio di Benchmark.

Il grafico mostra che all'aumentare della TEV cresce l'ampiezza dell'ellisse portandolo ad essere tangente alla frontiera di portafoglio efficiente. Un ulteriore aumento della TEV porta però ad uno spostamento dell'intero ellisse verso destra fino a divenire passante per il punto che rappresenta l'indice di riferimento. Tratto da Jorion (2003).

vincoli al *Value at Risk*³ (VaR)(Alexander and Baptista, 2008),(Palomba and Riccetti, 2012). Mentre Roll (1992), Jorion (2003) e Alexander and Baptista (2008) prendono in considerazione vincoli rispettivamente o sulla TEV o sul VaR, Palomba and Riccetti (2012) sviluppano un modello in cui possono essere prese in considerazione entrambi. All'interno della loro analisi vengono messi in luce due problemi da fronteggiare simultaneamente: il vincolo TEV potrebbe non essere compatibile con il vincolo su VaR ed i portafogli con vincoli sia sulla TEV, sia sul VaR sono spesso inefficienti. Viene per la prima volta

³Metodo di misurazione del rischio statistico che sintetizza la massima perdita attesa, su un orizzonte temporale prestabilito e nei limiti di un intervallo di confidenza. Esso viene definito come il θ -quantile della distribuzione di portafoglio, dove $0.5 < \theta < 1$; che rappresenta la minima perdita che sarà sostenuta con una probabilità di $1-\theta$

introdotto il concetto di *vincoli tollerabili* e viene mostrato come l'imposizione di un massimo alla TEV o al VaR permetterebbe al manager di spostarsi dalla *Constrained Mean-TEV Frontier* (Alexander and Baptista, 2008), quindi di selezionare un portafoglio con un rischio minore, riuscendo anche a ridurre la perdita di efficienza. L'intera analisi, oltre a presentare vari scenari che semplificano tutte le possibili interazioni tra le diverse frontiere di portafoglio (Markowitz (1952), Roll (1992), Alexander and Baptista (2008)) e a fornire soluzioni analitiche per tutte le intersezioni, introduce una nuova frontiera di portafoglio: la "*Fixed VaR-TEV Frontier*" (FVTF) (Palomba and Riccetti, 2012).

L'utilizzo dei primi due momenti delle distribuzioni comporta una perdita di informazione piuttosto rilevante ed una conseguente distorsione nell'utilità attesa dell'investitore. Autori come Arditti (1967) e Samuelson (1970) suggeriscono che una buona approssimazione della funzione di utilità attesa dell'investitore può essere ottenuta attraverso l'utilizzo di funzioni basate su momenti di ordine superiore al secondo associate a funzioni di distribuzioni dei rendimenti non-normali.

Jondeau and Rockinger (2005) dimostrano infatti che il criterio media-varianza, applicato su distribuzioni non-normali dei rendimenti, non risulta essere in grado di approssimare in maniera soddisfacente la funzione di utilità attesa dell'investitore.

All'interno della letteratura esistente sui modelli di asset allocation, costruiti sotto l'assunzione di non-normalità dei rendimenti, possiamo distinguere tre diversi filoni: il primo riguarda l'utilizzo di modelli regime switching (Ang and Bekaert (2002), Guidolin and Timmermann (2008)), il secondo comprende i modelli generalized conditional heteroskedasticity (GARCH) con distribuzioni delle innovazioni non normali, ed il terzo si sviluppa intorno allo strumento statistico delle *Copulae* (ad esempio Patton (2004), Jondeau and Rockinger (2006b), Riccetti (2013), De Lira Salvatierra and Patton (2015), Almeida, Czado, and Manner (2016)).

Numerosi studi empirici dimostrano la presenza di asimmetria negativa all'interno delle dinamiche di correlazione. Quest'ultima risulta infatti essere molto più forte nei periodi di crisi di mercato, piuttosto che in periodi di crescita di mercato. Nel tentativo di quantificare gli effetti di questa asimmetria all'interno di problemi di asset allocation, Ang and Bekaert (2002) sviluppano un modello che considera la presenza di dinamiche di tipo time-varying sia nella correlazione tra i diversi asset, sia nella volatilità dei loro rendimenti, applicando un modello *regime switching* in presenza di una funzione di utilità di tipo CRRA (Constant Relative Risk Aversion). Guidolin and Timmermann (2008) modificano i modelli standard di asset allocation, basati sul criterio media-varianza con l'assunzione di distribuzioni rendimenti invarianti nel tempo in due modi: inizialmente considerano una funzione di utilità degli investitori che dipende non solo dai primi due momenti, ma anche da momenti di ordine superiore. Successivamente costruiscono un modello in cui la distribuzione congiunta dei rendimenti dei vari asset segue un processo Markov switching, riuscendo così a catturare fenomeni di asimmetria e curtosi all'interno della distribuzione congiunta dei rendimenti.

Il secondo filone di letteratura, inerente all'uso di modelli GARCH con distribuzioni non normali delle innovazioni, comprende una larga parte della letteratura econometrica sull'allocation di portafoglio. Harvey, Liechty, Liechty, and Muller (2010) e Jondeau and Rockinger (2012) si soffermano sull'uso della distribuzione Normale asimmetrica introdotta da Sahu, Dey, and Branco (2003) mentre Patton (2004) e Jondeau and Rockinger (2006b) utilizzano la metodologia GARCH in un modello con le copulae.

Le distribuzioni non normali dei rendimenti sono state oggetto di studio anche in lavori non riguardanti problemi di asset allocation come ad esempio Longin and Solnik (2001) e Bauwens and Laurent (2005). Nel primo articolo gli autori dimostrano la non-normalità delle distribuzioni di rendimenti di asset

internazionali utilizzando una misura di *exceedance correlations*⁴ $\tilde{\rho}(q)$.

$$\tilde{\rho}(q) = \begin{cases} \text{corr}[X, Y | X > Q_x(q) \cap Y > Q_y(q)], & \text{per } q \geq 0.5 \\ \text{corr}[X, Y | X \leq Q_x(q) \cap Y \leq Q_y(q)], & \text{per } q \leq 0.5 \end{cases}$$

dove $Q_x(q)$ e $Q_y(q)$ sono i q -quantile di X ed Y . Bauwens and Laurent (2005), invece, nel loro studio costruiscono un modello GARCH multivariato considerando una distribuzione *skew-t* multivariata.

Un approccio comune per far fronte alla non-normalità dei rendimenti è quello di approssimare la funzione di utilità dell'investitore attraverso un'espansione in serie di Taylor per ordini superiori al secondo. Jondeau and Rockinger (2006b) sviluppano una metodologia di asset allocation considerando un'espansione fino al quarto ordine dei momenti campionari. In un successivo lavoro (Jondeau and Rockinger, 2012), ampliano il loro stesso modello utilizzando un *Dynamic Conditional Correlation* (DCC) GARCH. In entrambi questi lavori i residui vengono estratti da una distribuzione *skew-t* multivariata. Non avendo conoscenza, a priori, delle distribuzioni delle innovazioni, possono sorgere alcuni problemi di errata specificazione, quindi non è possibile stimare in maniera efficiente la matrice dei momenti. Kinoshita (2015), estendendo il lavoro di Jondeau and Rockinger (2006b), suggerisce di utilizzare un metodo di stima semi-parametrico per la stima della matrice dei momenti, associata ad un'approssimazione fino al quarto ordine della funzione di utilità dell'investitore.

Come già evidenziato, molti studi in campo finanziario fanno ampio uso della letteratura GARCH. Tali modelli, introdotti da Engle nel 1982 (modelli ARCH) e generalizzati successivamente da Bollerslev (1986), si concentrano sulla misurazione della volatilità assumendo che essa non sia osservabile e che sia in qualche modo legata al suo passato attraverso un processo di tipo auto-regressivo.

⁴Elevata probabilità che ad un elevato valore positivo/negativo di una variabile corrisponda un elevato valore positivo/negativo di un'altra variabile.

Hamao, Masulis, and Ng (1990), Susmel and Engle (1994), e Bekaert and Harvey (1995), nei loro rispettivi lavori, hanno analizzato l'interdipendenza tra rendimenti e volatilità nei diversi mercati azionari utilizzando alcune generalizzazioni multivariate di tali modelli. In studi più recenti, Kroner and Ng (1998), Engle and Sheppard (2001), Engle (2002) e Tse and Tsui (2002) sviluppano dei modelli GARCH con covarianze o correlazioni varianti nel tempo.

In presenza di distribuzioni dei rendimenti non normali, risulta difficoltoso costruire una distribuzione multivariata che legghi le diverse serie. Il terzo filone si concentra sull'utilizzo dello strumento matematico/statistico delle *Copulae* che permette di costruire una qualsiasi distribuzione multivariata avendo conoscenza solo delle distribuzioni marginali.

Ogni distribuzione di probabilità congiunta ha al suo interno informazioni riguardanti i comportamenti marginali delle diverse variabili, nonché la loro struttura di dipendenza. La Copula, conosciuta anche come funzione di dipendenza, ci permette di isolare la struttura di dipendenza, quindi di effettuare un'analisi separata della dipendenza e delle distribuzioni marginali.

Le Copulae sono state introdotte da Sklar (1959), e solo di recente si sono affermate nel mondo del risparmio gestito con applicazioni su problemi di costruzione di portafoglio e gestione del rischio.

I primi lavori in questo campo sono stati sviluppati da Patton (2004) e Jondeau and Rockinger (2006b). Risulta ben nota in letteratura la difficoltà di implementazione di tali modelli in ambito multivariato. Joe (1997), nel tentativo di ovviare a questo problema, introduce il metodo del *Pair Copula Construction* (PCC). Tale metodologia, utilizzando il concetto di fattorizzazione sequenziale, considera una batteria di copulae bivariate come una buona approssimazione di una copula multivariata.

Anche l'applicazione del PCC comporta diversi problemi, il primo fra tutti è l'aumento dei parametri da stimare con conseguente aumento dei tempi di stima. Per tale motivo la loro applicazione si è soffermata, per diversi anni, a studi che riguardavano un numero ristretto di variabili.

Nell'ultimo decennio la letteratura in ambito multivariato ha subito una crescita significativa (Brechmann, Czado, and Aas (2012), Almeida and Czado (2012), Brechmann and Czado (2013), Dißmann, Brechmann, Czado, and Kurovicka (2013)). Questi modelli, applicati su un numero piuttosto elevato di variabili, considerano però, come semplificazione, una struttura di dipendenza invariante nel tempo.

Come già evidenziato in precedenza, la letteratura econometrica contiene numerosi contributi che evidenziano come la volatilità condizionale delle serie storiche finanziarie sia soggetta a cambiamenti nel corso del tempo. Alcuni lavori hanno incentrato l'attenzione sull'utilizzo dello strumento statistico delle copulae al fine di far fronte a tale evidenza empirica. I primi lavori in merito si sono soffermati a delle strutture di dipendenza bivariate, come ad esempio Patton (2006) e Jondeau and Rockinger (2005). Studi più recenti hanno ampliato il campo di applicazione a problemi multivariati con un elevato numero di variabili (Almeida, Czado, and Manner, 2016). I lavori appena citati inseriscono una dinamica time-varying all'interno dei parametri della copula. In altre parole la forma funzionale della copula resta fissa mentre il/i parametro/i variano nel tempo seguendo come funzione dei valori ritardati del set informativo in maniera del tutto simile ai modelli ARCH. Altri lavori possono essere ricondotti alla letteratura *regime switching* (Hamilton, 1989). I modelli regime switching con copulae, a differenza di quanto visto nei modelli precedenti, spostano l'attenzione sulla forma funzionale della copulae. È infatti quest'ultima ad essere variante nel tempo e non più i parametri. Essi sono stati analizzati da diversi autori come Rodriguez (2007), Okimoto (2008), Chollete, Heinen, and Valdesogo (2009). Le copulae stocastiche sono state introdotte da Hafner and Manner (2012) e successivamente ampliate da Manner and Sergers (2011).

La Copula risulta essere lo strumento più flessibile di costruzione di funzioni di distribuzioni multivariate in quanto permette di legare tra loro variabili che possiedono funzioni di distribuzione marginali differenti. L'elevata variabilità delle serie dei rendimenti, nonché la presenza sempre più frequente di shock

finanziari, produce degli effetti anche all'interno della struttura di dipendenza. L'assunzione di una struttura di dipendenza costante nel tempo risulta pertanto essere estremamente semplificatrice della realtà. Per tale ragione questo lavoro si concentra sullo studio e sulla costruzione di copulae multivariate con il metodo PCC, che possano prendere in considerazione una struttura di dipendenza dinamica nel tempo. L'utilizzo della metodologia PCC permette di analizzare al meglio la relazione di dipendenza esistente tra ogni coppia di variabili e, sfruttando il meccanismo della fattorizzazione sequenziale, riesce a catturare l'intera struttura di dipendenza multivariata. L'inserimento di una struttura di dipendenza dinamica è dettata dal crescente bisogno di strumenti tecnici sempre più flessibili che possano cogliere al meglio le dinamiche temporali nel comportamento congiunto di un numero elevato di variabili finanziarie. Lo studio viene condotto sulla più generale e flessibile copula multivariata *Regular-Vine* (R-Vine) con all'interno una struttura di dipendenza dinamica guidata da una metodologia *Generalized Autoregressive Score* (GAS) introdotta da Creal, Koopman, and Lucas (2013).

Ad oggi gli studi condotti con copulae multivariate con dinamicità nei parametri si sono soffermati a strutture di tipo *Drawable-Vine* (D-vine) e *Canonical-Vine* (C-Vine). Esse sono dei casi particolari della R-vine e vincolano la costruzione di un modello multivariato su un percorso prestabilito (Vedi Paragrafo 2.1.4). Questo lavoro amplia la letteratura esistente inserendo una dinamica time-varying nei parametri di una copula R-vine.

Lo scopo è quello di introdurre un nuovo modello che possa essere utilizzato all'interno di tecniche di allocazione di portafoglio, andando ad indagare se l'utilizzo di un modello con copula multivariata dinamica possa avere una qualche utilità nella scelta dei pesi di un portafoglio azionario. Differentemente dal lavoro svolto da Riccetti (2013), in questo caso si è posta l'attenzione solo su titoli azionari appartenenti allo stesso mercato che sono tendenzialmente molto correlati tra loro. Si è preso in considerazione un periodo di forte incertezza ed instabilità del mercato italiano, dal 2010 al 2017, e lo studio viene

sviluppato su 17 serie di rendimenti di asset quotati all'interno dell'indice di borsa FTSE-MIB. La prima parte di questo lavoro si concentra sullo strumento statistico della Copula e sulla sua generalizzazione in ambito multivariato. I modelli con copula esistenti in letteratura fanno un ampio uso dei modelli GARCH (Bollerslev, 1986) nella trattazione della volatilità delle singole serie, pertanto, viene indagata brevemente la metodologia GARCH e la sua generalizzazione multivariata.

Nel terzo capitolo viene costruita una copula multivariata R-vine attraverso il metodo PCC con una struttura di dipendenza time-varying, vengono analizzate le performance in-sample e valutata la capacità previsiva confrontando il tutto con modelli con copula statici (R-Vine e D-vine) e due modelli DCC con innovazioni derivanti da una Normale e da una t .

Nel quarto ed ultimo capitolo viene proposto uno studio empirico su metodologie di asset allocation ed ottimizzazione di portafoglio, andando ad analizzare la rilevanza pratica per un portfolio manager. Parte integrante del lavoro riguarda l'implementazione dell'intera procedura di stima all'interno del programma econometrico *gretl*. I codici relativi alla stima di ciascuna copula bivariata utilizzata vengono riportati in Appendice. Parte di essi sono stati oggetto di presentazione all'interno della 5th *gretl* Conference. Al fine di validare i codici proposti, le stime sono state condotte anche con il programma statistico R ed i risultati ottenuti sono gli stessi.

Capitolo 2

Letteratura

Il concetto di volatilità all'interno dei mercati finanziari e l'interdipendenza esistente tra di essi, sono state le fonti dei principali studi econometrici degli ultimi decenni nel campo finanziario.

Questo lavoro si concentra sullo studio della dipendenza esistente tra diversi asset finanziari attraverso l'utilizzo delle Copulae. La letteratura in merito ha subito negli ultimi anni una crescita esponenziale passando dai più semplici modelli bivariati applicati su un numero ridotto di asset a modelli multivariati più flessibili e complessi che prendono in considerazione un numero elevato di variabili finanziarie. Vengono dapprima analizzati il concetto di Copula introdotto da Sklar (1959) ed i principali studi nel campo bivariato (Patton, 2004) per poi passare in seguito alla loro generalizzazione nel multivariato (*Pair Copula Construction* (Joe, 1997)) ed ai modelli sviluppati negli ultimi anni (Brechmann and Czado (2013), Dißmann, Brechmann, Czado, and Kurowicka (2013), De Lira Salvatierra and Patton (2015), Almeida, Czado, and Manner (2016)). Tutti questi modelli fanno un ampio uso dei modelli GARCH (Bollerslev, 1986) nella trattazione della volatilità delle singole serie, pertanto la parte finale di questo capitolo si sofferma sulla trattazione dei principali modelli di eteroschedasticità condizionale, introducendo i lavori riguardanti la loro generalizzazione nel multivariato.

2.1 Copulae

Le Copulae, introdotte da Sklar (1959), si sono di recente affermate nel mondo del risparmio gestito con applicazioni su problemi di costruzione di portafoglio e gestione del rischio. Data la loro grande flessibilità la letteratura in merito ha subito un rapido sviluppo.

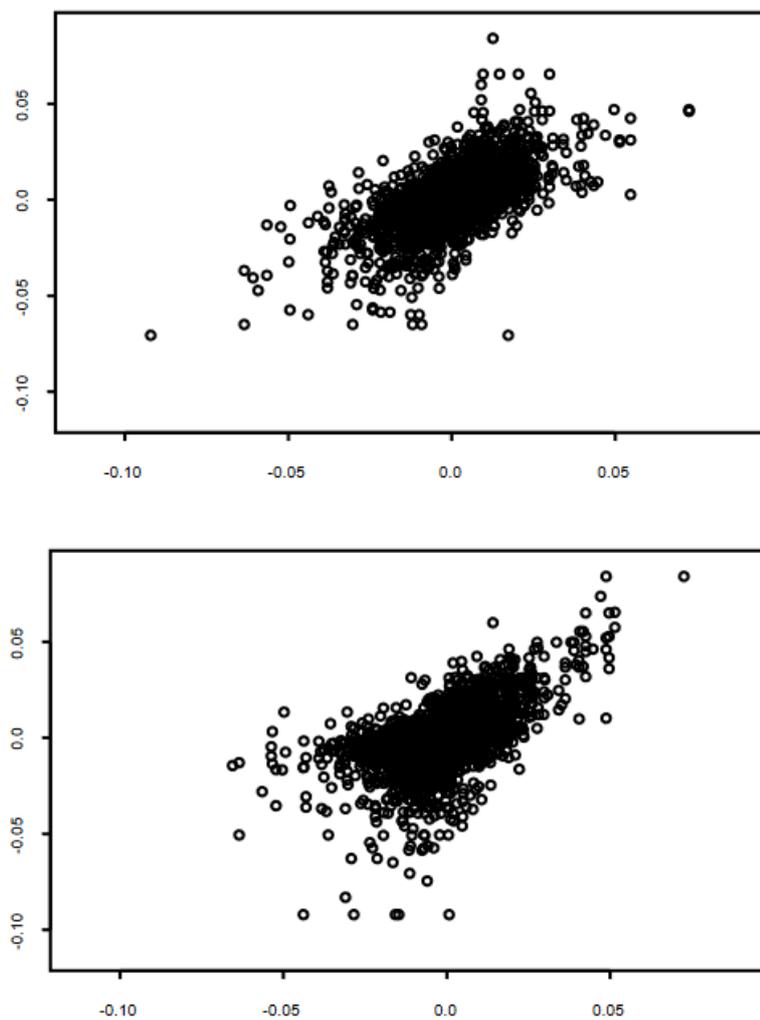
Di seguito vengono riportati i principali aspetti che caratterizzano le Copulae ed i lavori più rilevanti sviluppati negli ultimi anni.

Ogni distribuzione di probabilità congiunta ha al suo interno informazioni riguardanti i comportamenti marginali delle diverse variabili nonché la loro struttura di dipendenza. La copula permette di isolare la struttura di dipendenza multidimensionale, creando una funzione di distribuzione congiunta che consente di effettuare uno studio più approfondito rispetto alla semplice correlazione (Teorema 2.1). Una rappresentazione grafica è riportata nella Figura 2.1.1. Il grafico mostra come, prendendo in esame la semplice correlazione lineare si assume che il comportamento congiunto delle serie sia il medesimo, quando nella realtà sono presenti delle differenze.

Inizieremo introducendo il concetto di copula non-condizionata per poi passare alla più complessa, ma anche più utile per il nostro lavoro, copula condizionata.

Verranno di seguito indicate con le lettere maiuscole le funzioni di ripartizione, mentre con le minuscole le funzioni di densità. Come da convenzione le variabili casuali verranno riportate con lettere maiuscole X_t e le realizzazioni di tali variabili con lettere minuscole x_t .

Figura 2.1.1: *Strutture di dipendenza di funzioni di distribuzione bivariate*



Il grafico mostra la diversa struttura di dipendenza di due funzioni di distribuzioni bivariate che possiedono lo stesso coefficiente di correlazione.
Fonte: Aas (2004)

Ora assumiamo di non avere variabili condizionanti, quindi $X \sim F$, $Y \sim G$ e $(X, Y) \sim H$

Copula non-condizionata: la copula non condizionata di (X, Y) è la funzione di distribuzione congiunta di $U \equiv F(X)$ e $V \equiv G(Y)$ dove $X \sim F$, $Y \sim G$ e F e G sono continue.

Le variabili U e V sono conosciute come *probability integral transformation* di X e Y . Tali variabili si distribuiscono come una v.c. *Uniforme* $(0,1)$ (Fisher, 1932) .

Fisher (1932): sia X una variabile casuale con funzione di distribuzione continua F allora $U = F(x)$ si distribuisce come una v.c. *Uniforme* $(0,1)$, indipendentemente dalla funzione di distribuzione di partenza.

Per $0 \leq u \leq 1$ abbiamo:

$$\begin{aligned} Pr[U \leq u] &= Pr[F(X) \leq u] \\ &= Pr\{F^{(-1)}[F(X)] \leq F^{(-1)}(u)\} \\ &\equiv Pr[X \leq F^{(-1)}(u)] \\ &= u \end{aligned}$$

per $u < 0$ si ha $Pr[U \leq u] = 0$. Medesimo meccanismo per $u > 1$. Quindi $U \sim U(0, 1)$.

Questo enunciato sfrutta la definizione di funzione di ripartizione che risiede all'interno dell'intervallo $[0, 1]$ per costituzione.

Abbiamo bisogno di due proprietà per permettere alla copula di rispettare le caratteristiche proprie delle funzioni di distribuzione.

Consistenza: sia $Z : \mathfrak{R}^n \rightarrow \mathfrak{R}$ con dominio $Z = [D_1 \times \dots \times D_n]$ dove d_i è il valore più piccolo degli insiemi D_i non vuoti.

La funzione è detta consistente se si annulla per un generico valore $g_k \subset \mathfrak{R}^n$

tale che $g_k = a_k$

$$Z(g) = Z(g_1, \dots, g_{k-1}, a_k, g_{k+1}, \dots, g_k) = 0 \quad (2.1.0.1)$$

In pratica, una funzione è consistente, se si annulla quando uno solo degli insiemi D_i assume il suo valore più piccolo nel suo dominio.

Per spiegare la proprietà di n - *crescita* bisogna partire da un generico n - *intervallo* definito come prodotto cartesiano di n intervalli chiusi: $A = [a_{11}, a_{12}] \times \dots \times [a_{n1}, a_{n2}]$ con $a_{i1} \leq a_{i2}$, $i = 1, 2, \dots, n$. Indichiamo il fattore c come un generico vertice di A e con $ver(A)$ l'insieme di tutti i vertici di A : allora $c \in ver(A)$ se e solo se la sua i -esima componente w_i , con $i = 1, \dots, n$ è uguale a a_{i1} o a_{i2} . Il primo termine della disequazione viene così definito:

$$\prod_{i=1}^n \text{sgn}(2c_i - a_{i1} - a_{i2}) = \begin{cases} -1 & \text{se } a_{i1} \neq a_{i2}, \text{ per } i \text{ dispari} \\ 0 & \text{se } \exists i : a_{i1} = a_{i2} \\ 1 & \text{se } a_{i1} \neq a_{i2}, \text{ per } i \text{ pari} \end{cases} \quad (2.1.0.2)$$

Se $ver(A) \subset DomZ$ il Z - *volume* di A è rappresentato come

$$\sum_{c \in ver(A)} Z(c) \prod_{i=1}^n \text{sgn}(2c_i - a_{i1} - a_{i2})$$

che risulta essere di fondamentale importanza per il concetto di n -*crescenza*.

Crescita: la funzione Z è definita ***n-crescente*** se il Z -*volume* di D risulta non negativo per ogni n -*scatola* di D ($A = [a_{i1}, a_{i2}] \times \dots \times [a_{n1}, a_{n2}]$) e $ver(A) \subset DomZ$:

$$\sum_{c \in ver(A)} Z(c) \prod_{i=1}^n \text{sgn}(2c_i - a_{i1} - a_{i2}) \geq 0 \quad (2.1.0.3)$$

dove sgn rappresenta la funzione segno, $DomZ$ il dominio della funzione Z e $ver(A)$ l'insieme dei vertici di A .

Copula bidimensionale: di seguito vengono elencate 4 caratteristiche delle copulae bidimensionali:

1. $C : [0, 1] \times [0, 1] \rightarrow [0, 1]$
2. $C(u, 0) = C(v, 0) = 0$ per $u, v \in [0, 1]$, oppure $C(0, u) = C(0, v) = 0$
3. $C(u, 1) = u$ e $C(1, v) = v$ per $u, v \in [0, 1]$
4. $V_C[u_1, u_2] \times [v_1, v_2] \equiv C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$
 $\forall u_1, u_2, v_1, v_2 \in [0, 1]$ dove $u_1 \leq u_2$ e $v_1 \leq v_2$.

dove per V_C viene inteso il volume. Queste proprietà si estendono anche nel caso multidimensionale.

Il Teorema di Sklar (1959) è il prodotto più rilevante di tutta la teoria delle copulae. Egli infatti ci fornisce una tecnica che consente di costruire distribuzioni multivariate con marginali e copulae arbitrarie.

Teorema di Sklar 2.1 *Sia F una funzione di distribuzione congiunta con marginali F_1, \dots, F_n . Esiste allora una copula $C : [0, 1]^n \rightarrow [0, 1]$ per ogni x_1, \dots, x_n definite in $\mathfrak{R} = (-\infty, +\infty)$ tale che:*

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$$

Se le marginali sono continue, allora la copula C è unica altrimenti è unicamente determinata in $Rk(F_1) \times \dots \times Rk(F_n)$ (dove Rk indica il rango). Al contrario, se C è la copula e F_1, \dots, F_n sono funzioni di distribuzione marginali univariate, allora F è la loro funzione di distribuzione congiunta.

$$C(u_1, \dots, u_n) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)).$$

Possiamo ricavare la probabilità congiunta cumulata in funzione delle marginali.

Il Teorema 2.1 permette di tirare fuori dalle funzioni di distribuzione multivariate la struttura di dipendenza, rappresentata dalla copula, e le marginali univariate.

Per comprendere a pieno il Teorema 2.1 c'è bisogno di definire la funzione inversa F^{-1} . Essa è la funzione inversa di una data funzione di distribuzione cumulata univariata:

$$F^{-1}(m) = \inf \{x \in \mathbb{R} \mid F(x) \geq m\}.$$

Copula condizionata: la copula condizionata di $(X, Y)|W$, con $X|W \sim F$ e $Y|W \sim G$, è la funzione di distribuzione congiunta condizionata di $U \equiv F(X|W)$ e $V \equiv G(Y|W)$ data W .

Come per le copulae non condizionate anche queste seguono alcune proprietà:

1. $C(u, 0|w) = C(0, v|w) = 0$ per $(u, v) \in [0, 1]$;
2. $C(u, 1) = u$ e $C(1, v) = v$ per $u, v \in [0, 1]$
3. $V_C([u_1, u_2] \times [v_1, v_2]|w) \equiv C(u_2, v_2|w) - C(u_2, v_1|w) - C(u_1, v_2|w) + C(u_1, v_1|w) \geq 0$ per $u_1, u_2, v_1, v_2 \in [0, 1]$ dove $u_1 \leq u_2$ e $v_1 \leq v_2$.

Il teorema 2.1 può essere facilmente esteso anche in questo caso nel seguente modo:

$$H(x, y|w) = C(F(x|w), G(y|w)|w), \forall (x, y) \in \mathfrak{R} \times \mathfrak{R} \text{ e } w \in \mathcal{W} \quad (2.1.0.4)$$

La funzione di densità relativa alla (2.1.0.4) può essere facilmente ricavata,

purché F e G siano differenziabili, e H e C siano differenziabili due volte:

$$\begin{aligned}
 h(x, y|w) &\equiv \frac{\partial^2 H(x, y|w)}{\partial x \partial y} \\
 &= \frac{\partial F(x|w)}{\partial x} \cdot \frac{\partial G(y|w)}{\partial y} \cdot \frac{\partial^2 C(F(x|w), G(y|w)|w)}{\partial u \partial v} \\
 h(x, y|w) &\equiv f(x|w) \cdot g(y|w) \cdot c(u, v|w), \tag{2.1.0.5} \\
 &\text{per } (x, y, w) \in \mathfrak{R} \times \mathfrak{R} \times \mathcal{W}
 \end{aligned}$$

$$\mathcal{L}_{XY} = \mathcal{L}_X + \mathcal{L}_Y + \mathcal{L}_C \tag{2.1.0.6}$$

dove \mathcal{W} rappresenta il supporto di W , $u \equiv F(x|w)$, $v \equiv G(y|w)$, $\mathcal{L}_{XY} \equiv \log h(x, y|w)$, $\mathcal{L}_X \equiv \log f(x|w)$, $\mathcal{L}_Y \equiv \log g(y|w)$ e $\mathcal{L}_C \equiv \log c(u, v|w)$.

La definizione appena riportata mostra qualche complicazione data dal fatto che il set informativo condizionante deve essere lo stesso per tutte le marginali e per la copula stessa. Il mancato utilizzo dello stesso set informativo comporta che l'equazione (2.1.0.4) non rappresenti più una valida funzione di distribuzione.

L'intera famiglia delle copulae gode di diverse proprietà:

- Invarianza rispetto a trasformazioni delle marginali strettamente crescenti (Schweizer e Wolff, 1976). Se $\alpha_1, \dots, \alpha_d$ sono trasformazioni monotone crescenti in $\alpha_i : \mathfrak{R}(F_i) \rightarrow \mathfrak{R}^d$, le variabili casuali $\alpha_1(X_1), \dots, \alpha_d(X_d)$ con marginali

$$G_1 = F_1(\alpha_1^{-1}), \dots, G_d = F_d(\alpha_d^{-1})$$

hanno la copula

$$C_{\alpha_1(X_1), \dots, \alpha_d(X_d)}(u) = C_{x_1, \dots, x_d}(u)$$

- Rotazione: $\bar{U} = 1 - u$ e $\bar{V} = 1 - v$ sono entrambe delle variabili casuali uniformi standardizzate che seguono le seguenti regole:

$$- \bar{U}, \bar{V} \text{ hanno copula } C^{-+}(u, v) = v - C(1 - u, v);$$

- U, \bar{V} hanno copula $C^{+-}(u, v) = u - C(u, 1 - v)$;
- \bar{U}, \bar{V} hanno copula $C^{--}(u, v) = u + v - 1 + C(1 - u, 1 - v)$

- Limiti: per ogni copula $C(u_1, \dots, u_d)$ esistono un limite superiore ed uno inferiore conosciuti meglio come limiti di Fréchet:

$$\max \sum_{i=1}^d (u_i + 1 - d, 0) \leq C(u) \leq \min(u_1, \dots, u_d) \quad (2.1.0.7)$$

dove d è il massimo numero di elementi.

Un concetto di fondamentale importanza nell'analisi della dipendenza tra valori estremi è la *dipendenza sulle code* (*Tail Dependency*). L'attenzione in questo caso si concentra sulle code della funzione di distribuzione congiunta. Essendo tale concetto una proprietà propria delle copulae è anch'esso invariante rispetto a trasformazioni crescenti delle variabili casuali.

La dipendenza sulla coda superiore (λ_u) è la probabilità che X_n sia maggiore del suo q -quantile ($F(x = q) = 1 - u$), dato X_{n-1} maggiore del suo q -quantile, con q tendente al proprio limite:

$$\lambda_u = \lim_{u \rightarrow 0^+} P(F_2(x_2) > 1 - u | F_1(x_1) > 1 - u) \quad (2.1.0.8)$$

dove $\lambda_u \in [0, 1]$. Se $\lambda_u \in (0, 1]$ saremo sicuramente in presenza di dipendenza in coda, mentre nel caso in cui $\lambda_u = 0$ allora le due variabili sarebbero asintoticamente indipendenti. In maniera analoga si ricava la dipendenza sulla coda inferiore $\lambda_l \in [0, 1]$ data dalla relazione:

$$\lambda_l = \lim_{u \rightarrow 1^-} P(F_2(x_2) < 1 - u | F_1(x_1) < 1 - u) \quad (2.1.0.9)$$

I coefficienti λ_l e λ_u misurano la probabilità che due diverse variabili assumano entrambe valori estremi. Come la leptocurtosi nelle distribuzioni di singole variabili casuali è indice di una più elevata probabilità di assistere ad eventi estremi, la presenza di dipendenza sulla coda diversa da zero aumenta la

probabilità di avere eventi estremi comuni. La grande attenzione che il risk manager pone su strumenti come il Value-at-Risk (VaR) e altri strumenti utili per stimare la probabilità di incorrere in grandi perdite, rende la presenza della dipendenza non-normale di grande interesse. Cherubini and Luciano (2001) e Embrechts and Juri (2003) analizzano il VaR di portafoglio utilizzando il metodo delle copulae. Rosenberg and Schuermann (2006), McNeil and Embrechts (2005) e Alexander (2008) utilizzano la copula nell'ambito di problemi di risk management.

Per una generalizzazione multivariata della dipendenza in coda appena definita si veda Schmidt and Stadtmuller (2003).

2.1.1 Famiglie di Copulae

Esistono in letteratura due tipologie di copulae:

- esplicite: possiedono un'espressione in forma chiusa e seguono le generali metodologie matematiche utilizzate per costruirle (copula gaussiana (2.1.1.2) e t (2.1.1.3));
- implicite: sono ricavate da funzioni di distribuzione multivariate conosciute, ma non necessariamente possiedono un'espressione in forma chiusa (copulae di Archimede o Archimedée).

Le copulae Archimedée hanno il vantaggio di possedere pochi parametri nella rappresentazione della struttura della dipendenza, ma sono poco flessibili, ovvero hanno una scarsa capacità di adattamento ai dati, soprattutto in presenza di un numero elevato di variabili. Tra le principali troviamo la copula Gumbel, la Clayton e la Frank. Esse derivano tutte dalle seguente espressione:

$$C(u_1, u_2) = \phi^{-1}[\phi(u_1) + \phi(u_2)] \quad (2.1.1.1)$$

dove ϕ è una funzione decrescente, conosciuta come *generatore* che soddisfa le due seguenti condizioni: $\phi(0) = \infty$ e $\phi(1) = 0$.

Ognuna di queste copula possiede un diverso *generatore*. Alcuni esempi sono riportati in Tabella 1.1.

Tabella 2.1.1: *Generatori e valori ammissibili per il parametro θ*

Copula	generatore	valori
Gumbel	$(-\ln t)^\theta$	$\theta \geq 1$
Clayton	$\frac{1}{\theta}(t^{-\theta} - 1)$	$\theta \geq -1$
Frank	$-\ln\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right)$	$\theta \in \Re$

In questo lavoro prenderemo in considerazione 7 diverse tipologie di Copulae (Gaussiana, t di Student, Gumbel, Rotated-Gumbel (180°), Clayton, Rotated-Clayton (180°), Frank).

1 . Copula Gaussiana:

$$c_N(u_1, u_2, \rho) = \frac{1}{\sqrt{1-\rho^2}} \exp\left(-\frac{\rho^2(\Phi^{-1}(u_1))^2 + \Phi^{-1}(u_2))^2 - 2\rho\Phi^{-1}(u_1)\Phi^{-1}(u_2)}{2(1-\rho^2)}\right) \quad (2.1.1.2)$$

dove Φ^{-1} rappresenta l'inversa della funzione di distribuzione cumulata della normale standardizzata univariata e ρ l'indice di correlazione lineare tra u_1 e u_2 .

La dipendenza sulla code per la copula Gaussiana è nulla da entrambe i lati:

$$\lambda_u = 0 \quad \lambda_l = 0$$

2 . Copula t-Student:

$$c_T(u_1, u_2, \rho, n) = \frac{\Gamma\left(\frac{n+2}{2}\right)\Gamma\left(\frac{n}{2}\right)\left(1 + \frac{\psi'\Omega\psi}{n}\right)^{-\frac{n+2}{2}}}{\sqrt{1-\rho^2}\left(\Gamma\left(\frac{n+1}{2}\right)\right)^2 \prod_{i=1}^2 \left(1 + \frac{\psi_i^2}{n}\right)^{-\frac{(n+1)}{2}}} \quad (2.1.1.3)$$

dove $\psi^{-1} = (t_n^{-1}(u_1), t_n^{-1}(u_2))$, t_n^{-1} rappresenta l'inversa della funzione di distribuzione cumulata della t-Student univariata con v gradi di libertà, ρ l'indice

di correlazione e Ω è la matrice di correlazione, mentre $\Gamma(\cdot)$ rappresenta la funzione gamma.

In questo caso, a differenza della copula Gaussiana, siamo in presenza di un indice di dipendenza sulle code non nullo e simmetrico:

$$\begin{aligned}\lambda_u &= 2t_{n+1} \left(-\sqrt{\frac{(n+1)(1-\rho)}{1+\rho}} \right) \\ \lambda_l &= 2t_{n+1} \left(-\sqrt{\frac{(n+1)(1-\rho)}{1+\rho}} \right)\end{aligned}\quad (2.1.1.4)$$

3. Copula Frank:

$$c_F(u_1, u_2, \theta) = \frac{\theta(1 - e^{-\theta})e^{-\theta(u_1+u_2)}}{(1 - e^{-\theta}) - (1 - e^{-\theta u_1})(1 - e^{-\theta u_2})}\quad (2.1.1.5)$$

con

$$\lambda_u = 0 \quad \lambda_l = 0$$

4. Copula Clayton:

$$c_C(u_1, u_2, \theta) = (1 + \theta)(u_1 u_2)^{-\theta-1} (u_1^{-\theta} + u_2^{-\theta} - 1)^{-2-1/\theta}\quad (2.1.1.6)$$

La Copula Clayton presenta un eccesso di dipendenza negativa (alta probabilità che ad un elevato valore negativo di una variabile corrisponde un elevato valore negativo dell'altra):

$$\lambda_u = 0 \quad \lambda_l = \begin{cases} 2^{-1/\theta} & \theta > 0 \\ 0 & \theta \leq 0 \end{cases}\quad (2.1.1.7)$$

5. Copula Gumbel:

$$c_G(u_1, u_2, \theta) = \frac{\exp[-((-\ln u_1)^\theta + (-\ln u_2)^{-\theta})^{1/\theta}](\ln u_1 \ln u_2)^{\theta-1}}{u_1 u_2 ((-\ln u_1)^\theta + (-\ln u_2)^\theta)^{2-1/\theta}} \times \quad (2.1.1.8)$$

$$\times \left[((-\ln u_1)^\theta + (-\ln u_2)^\theta)^{1/\theta} + \theta - 1 \right]$$

In questo caso, contrariamente a quanto avviene per la Clayton, l'eccesso di dipendenza si concentra solo sulla coda destra:

$$\lambda_u = 2 - 2^{1/\theta} \quad \lambda_l = 0 \quad (2.1.1.9)$$

6. 7. Rotated-Clayton (180°), Rotated-Gumbel (180°).

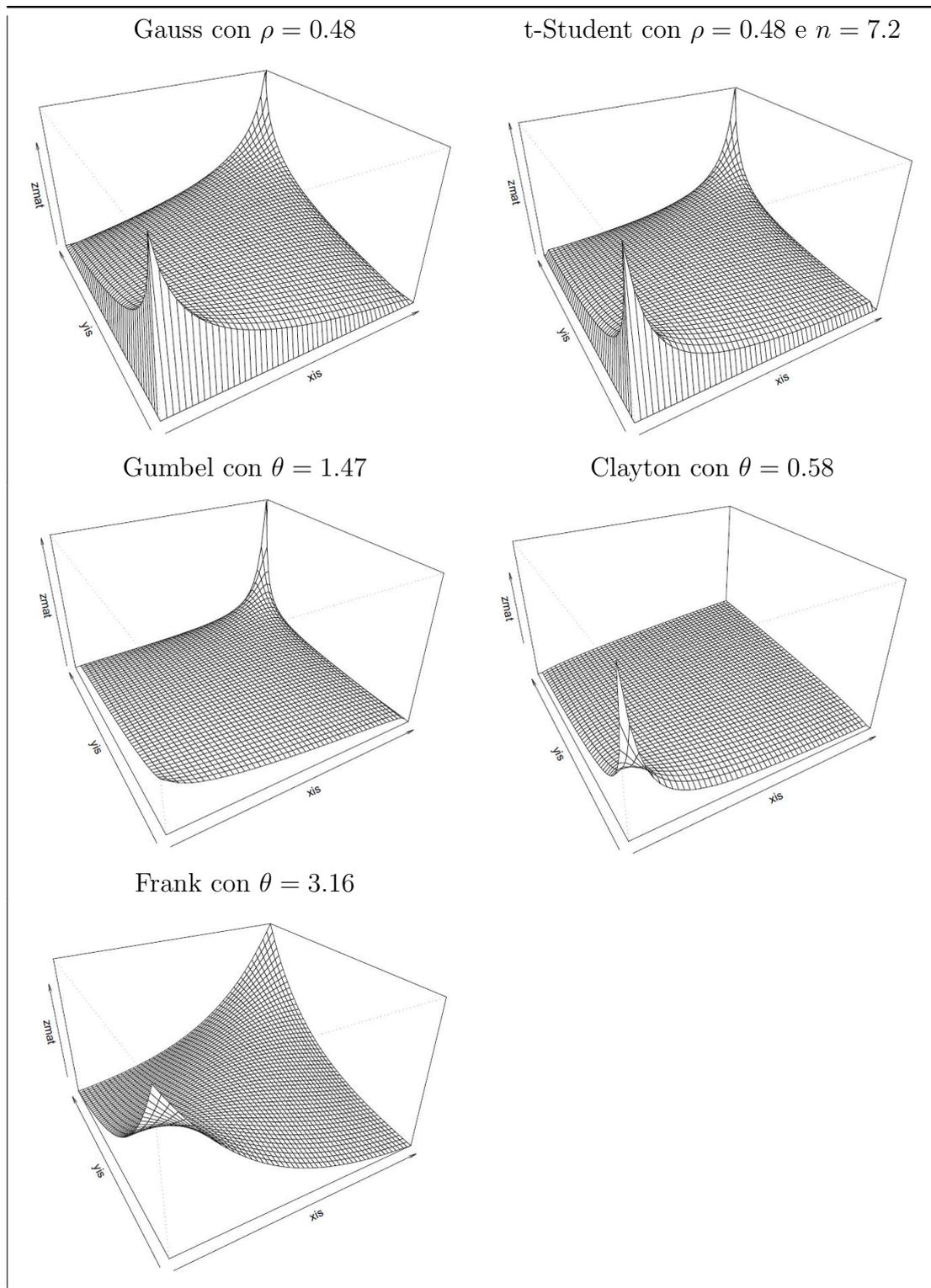
Le espressioni analitiche di queste due copulae, anche conosciute come Survival Copula, sono le medesime delle equazioni (2.1.1.6) e (2.1.1.8) dove però u_1 ed u_2 vengono sostituiti da:

$$\begin{cases} 1 - u_1 \\ 1 - u_2 \end{cases}$$

Gli indici di dipendenza sulle code, derivanti da questa trasformazione, risultano essere invertiti rispetto alle equazioni (2.1.1.7) e (2.1.1.9). Avremo infatti un eccesso di dipendenza sulla coda destra per la copula Rotated-Clayton ed un eccesso di dipendenza sulla coda sinistra per la Rotated-Gumbel. I valori numerici dei due indici sono gli stessi delle versioni non ruotate. Nella Figura 2.1.2 vengono rappresentate le funzioni di densità costruite utilizzando il Teorema 2.1.

All'interno di applicazioni finanziarie la dipendenza sulle code comporta alcune importanti implicazioni. Ad esempio, se la copula con cui sono modellate due diverse serie storiche di rendimenti è una t-Student piuttosto che una Gaussiana, la probabilità che i rendimenti delle due attività si trovino entrambe su una delle due code risulta essere più elevata.

Figura 2.1.2: Densità delle diverse copulae



a

^aFonte Riccetti(2010)

2.1.2 Stima delle Copulae

In questo capitolo vengono presi in esame due metodi di stima dei modelli con copula: parametrica e semi-parametrica. Entrambi i metodi fanno ricorso al stima di Massima Verosimiglianza.

Attraverso il Teorema 2.1 una qualsiasi funzione di densità f n -dimensionale avente marginali f_1, f_2, \dots, f_n può essere riscritta come:

$$f(x_1, x_2, \dots, x_n) = c(F_1(x_1), F_2(x_2), \dots, F_n(x_n)) \prod_{i=1}^n f_i(x_i)$$

dove c è la funzione di densità di una copula n -dimensionale.

Da tale rappresentazione possiamo scrivere la seguente funzione di Verosimiglianza:

$$L = \sum_{i=1}^n L_i + L_C \quad (2.1.2.1)$$

con $L_C = \sum_{k=1}^n \log c(F_1(x_1^{(k)}), F_2(x_2^{(k)}), \dots, F_n(x_n^{(k)}))$ che indica il contributo alla log-verosimiglianza della copula n -dimensionale e $L_i = \sum_{k=1}^n \log f_i(x_i^{(k)})$, $i = 1, 2, \dots, n$

Poniamo θ il vettore dei parametri della copula e α_i il vettore dei parametri delle distribuzioni marginali $f_i = f_i(x_i; \alpha_i)$. La funzione di log-verosimiglianza può essere scritta come:

$$l(\Theta) = \sum_{i=1}^n l_i(\alpha_i) + l_C(\theta)$$

La stima di massima verosimiglianza consiste nella massimizzazione della funzione di log-verosimiglianza all'interno del vettore di parametri $\Theta = (\alpha_i, \theta)$.

$$\Theta_{ML} = \operatorname{argmax}_{\theta} l(\Theta) \quad (2.1.2.2)$$

Tale procedura, conosciuta come *Full Maximal Likelihood* è però scarsamente utilizzata perché comporta dei tempi di stima molto elevati, in quanto debbo-

no essere stimati simultaneamente i parametri delle marginali e della copula. Per superare questo limite Shin and Louis (1995) e Joe and Xu (1996) introducono la procedura di stima a due stadi (*two-step*) che consiste nello stimare separatamente i parametri delle marginali e della copula attraverso il metodo della massima verosimiglianza. Si inizia dapprima con la stima dei parametri delle distribuzioni marginali per poi passare successivamente alla stima della struttura di dipendenza.

La procedura di stima a due passi parametrica utilizza il metodo *Inference Function of Margins* (IFM) (Joe, 1997). Nel primo step della procedura di inferenza vengono stimati i parametri delle marginali con il metodo della Massima Verosimiglianza, $\hat{\alpha}_i = \operatorname{argmax}_{\alpha_i} L_i(\alpha_i)$. Per questa prima parte è necessario conoscere le funzioni di distribuzione delle diverse variabili considerate.

La stima del parametro $\hat{\theta}$ viene effettuata attraverso la massimizzazione della funzione di log-verosimiglianza l_C utilizzando le stime dei parametri $\hat{\alpha}_i$ all'interno delle distribuzioni marginali.

$$\hat{\theta} = \operatorname{argmax} l_C(F_1(x_1; \hat{\alpha}_1), F_2(x_2; \hat{\alpha}_2), \dots, F_n(x_n; \hat{\alpha}_n))$$

Sotto le condizioni standard di regolarità le stime ottenute con il metodo IFM sono consistenti ed asintoticamente normali (Serfling, 1980). La stima a più step è però asintoticamente meno efficiente della *Full Maximum Likelihood*. Joe (1997) e Patton (2006) dimostrano che tale perdita di efficienza, nella maggior parte dei casi, non è statisticamente rilevante. La stima della matrice di covarianza dei parametri della funzione log-verosimiglianza risulta essere molto complessa, sia analiticamente, sia numericamente, in quanto richiede il calcolo di numerose derivate, pertanto, vengono utilizzati metodi di ricampionamento come il *jackknife* (Quenouille, 1949) ed altri metodi ad esso strettamente collegati. Sotto le classiche condizioni di regolarità (White, 1994), il metodo di stima multi step è asintoticamente normale ma, in questo caso, la matrice di covarianza asintotica assume una forma non-standard (dettagli sulla sua stima

in Patton (2006)).

Joe (1997) dimostra nel suo lavoro, attraverso delle simulazioni di Monte Carlo che l'approccio IFM è un'alternativa efficiente alla Full-MLE nella stima dei parametri di modelli multivariati.

Seguire un approccio di tipo parametrico può comportare problemi di errata specificazione in quanto le distribuzioni marginali delle diverse variabili potrebbero essere non note.

Il metodo di stima semi-parametrico permette di ovviare a questa problematica utilizzando una stima delle funzioni di distribuzione. Anche questa procedura, come la precedente, segue una metodologia a due passi. Inizialmente viene utilizzata l'*empirical distribution function* come stima delle funzioni di distribuzione marginali e successivamente viene massimizzato il contributo della log-verosimiglianza della copula l_C per la stima del parametro θ :

$$\hat{\theta} = \operatorname{argmax}_{\theta} l_C(\hat{F}_1(x_1), \hat{F}_2(x_2), \dots, \hat{F}_n(x_n)) \quad (2.1.2.3)$$

dove le funzioni \hat{F} rappresentano le funzioni di distribuzioni empiriche.

Le proprietà di tale stimatore e la stima della matrice di covarianza asintotiche dei parametri verranno descritte più avanti all'interno del paragrafo 2.1.5. Genest, Ghoudi, and Rivest (1993) dimostrano che, date le assunzioni standard di regolarità, lo stimatore $\hat{\theta}$ è uno stimatore consistente ed asintoticamente normale del parametro di dipendenza θ . Chen and Fan (2006) ricavano le condizioni standard sotto la quale ottenere una distribuzione asintotica normale e forniscono anche un metodo per la stima della matrice di covarianza asintotica.

2.1.3 Pair Copula Construction

Mentre nel campo bivariato esistono svariate tipologie di copulae più o meno flessibili (paragrafo 2.1.1), nel multivariato, sfortunatamente, la scelta risulta

essere molto limitata. Alla base di ciò troviamo sia la difficoltà nell'implementazione di copulae in grado di descrivere la struttura di dipendenza esistente tra un numero elevato di variabili, sia la loro scarsa capacità di adattamento ai dati. Ad esempio la copula Gaussiana non prende in considerazione la possibile presenza di una dipendenza in coda (*Tail Dependence*), mentre la copula t-Student è in grado di catturare solo episodi di dipendenza in coda simmetrici. Le copulae Archimedee (nel nostro caso Gumbel, Clayton e Frank) permettono lo studio di un tipo di dipendenza asimmetrica ma presentano un solo parametro per la sua descrizione e ciò è chiaramente troppo restrittivo quando si prendono in considerazione modelli con diverse variabili.

Il metodo *Pair Copula Construction* (d'ora in avanti PCC) introdotto da Joe (1996) e successivamente approfondito da Bedford and Cooke (2001), Bedford and Cooke (2002) permette il superamento di questi limiti.

L'idea è che una generale copula multivariata possa essere scomposta in una serie di copulae bivariate.

Partiamo dalla semplice definizione di funzione di densità congiunta utilizzando le funzioni di densità condizionali (fattorizzazione sequenziale):

$$f(y_1, \dots, y_n) = f(y_1) \cdot f(y_2|y_1) \cdot f(y_3|y_1, y_2) \cdot f(y_n|y_1, \dots, y_{n-1})$$

Scomponendo mediante una copula condizionale otteniamo:

$$\begin{aligned} f(y_2|y_1) &= \frac{f(y_1, y_2)}{f(y_1)} \\ &= c_{12}(F(y_1), F(y_2)) \cdot f(y_2) \end{aligned} \quad (2.1.3.1)$$

$$\begin{aligned} f(y_3|y_1, y_2) &= \frac{f(y_2, y_3|y_1)}{f(y_2|y_1)} \\ &= c_{23|1}(F(y_2|y_1), F(y_3|y_1)) \cdot f(y_3|y_1) \end{aligned} \quad (2.1.3.2)$$

combinando le precedenti scomposizioni si arriva, nel caso trivariato, al

seguinte risultato:

$$f(y_1, y_2, y_3) = c_{23|1}(F_{2|1}(y_2|y_1), F_{3|1}(y_3|y_1))c_{12}(F_1(y_1), F_2(y_2)), \\ c_{13}(F_1(y_1), F_3(y_3))f_1(y_1)f_2(y_2)f_3(y_3) \quad (2.1.3.3)$$

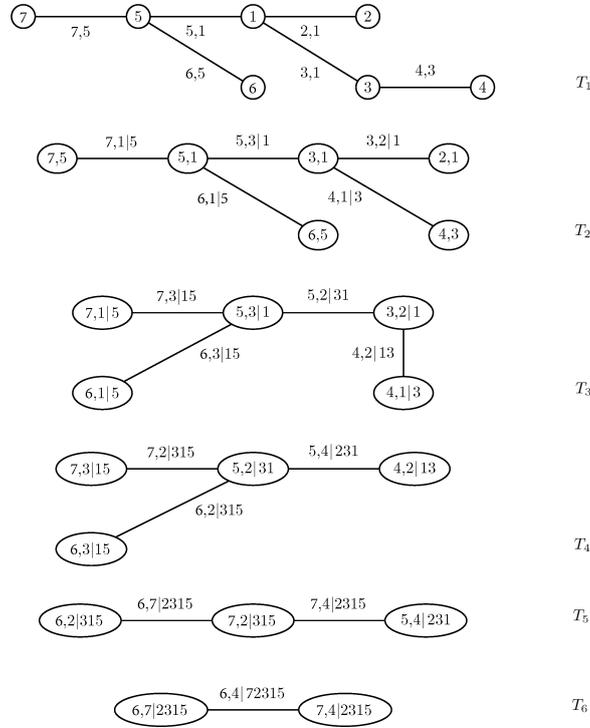
L'equazione (2.1.3.3) rappresenta una funzione di densità congiunta trivariata in termini di copulae bivariate C_{12} , C_{13} , $C_{23/1}$. Queste ultime possono essere scelte in maniera indipendente dalle altre, riuscendo così a modellare la diversa struttura di dipendenza esistente all'interno di un campione di variabili molto ampio. Le copulae costruite attraverso l'utilizzo del PCC sono anche conosciute come *Vine Copula*. Nel prossimo paragrafo verrà descritta nel dettaglio la copula *Regular vine* che sarà l'oggetto principale di questo lavoro, rivolgendo particolare attenzione alla procedura di stima ed ai problemi di inferenza (Dißmann, Brechmann, Czado, and Kurowicka, 2013), (Hobæk Haff, 2013).

2.1.4 R-vine copula

La copula *Regular-vine* (R-vine) rappresenta la forma più generale e più flessibile tra le vine copula. In letteratura sono stati utilizzati ampiamente due sottomodelli della stessa: la *Canonical-vine* (C-vine) e la *Drawable-vine* (D-vine). Queste tipologie sono state trattate da numerosi autori in diversi articoli tra cui Bedford and Cooke (2001), Bedford and Cooke (2002), Kurowicka and Cooke (2006) e Heinen and Valdesogo (2008). Cherubini, Luciano, and Vecchiato (2004), Aas, Czado, Frigessi, and Bakken (2007) e Berg and Aas (2007) sono stati i primi ad utilizzarle in applicazioni finanziarie.

Un modo chiaro ed intuitivo di rappresentazione delle vine copulae è attraverso uno schema ad albero sequenziale (Bedford and Cooke, 2001). Nella Figura 2.1.2 $T_i = (T_1, \dots, T_{n-1})$ rappresentano i vari tree della scomposizione, i numeri all'interno dei cerchi vengono chiamati *nodi* e costituiscono le distribuzioni marginali delle singole variabili prese in considerazione ed i numeri corrispondenti alle congiunzioni (*edge*) rappresentano le coppie di variabili sul-

Figura 2.1.3: R-vine copula con 7 variabili



Il grafico rappresenta un esempio di scomposizione R-vine con 7 variabili

le quali vengono applicate le copulae bivariate. Per $i = 2, \dots, n - 1$ vengono considerate le copulae bivariate condizionate.

Di seguito viene fornita la definizione di una copula *R-vine*.

R-vine 2.1 $V = (T_1, \dots, T_{n-1})$ è una R-vine con n elementi se:

1. T_1 è un tree con $N_1 = (1, \dots, n)$ nodi e E_1 l'insieme delle edge.
2. Per $i = 2, \dots, n - 1$, T_i è un tree con $N_i = E_{i-1}$ nodi ed insieme delle edge E_i .
3. Per $i = 2, \dots, n - 1$ e $(a, b) \in E_i$ con $a = (a_1, a_2)$ e $b = (b_1, b_2)$ deve valere che $\#(a \cap b) = 1$ (condizione di prossimità).

La copula *regular vine* costruita su un numero n di variabili consiste in una sequenza di diversi tree T_1, \dots, T_{n-1} con N_i nodi ed E_i edge per $i = 1, \dots, n - 1$,

dove T_1 abbiamo $N_1 = \{1, \dots, n\}$ nodi ed E_1 teste, e per $i = 2, \dots, n - 1$ l' i -esimo tree (T_i) ha come nodi $N_i = E_i$.

Per comprendere meglio tale definizione prendiamo ad esempio il modello rappresentato nella Figura 2.1.2 dove vengono stimate in T_1 , 6 copulae non condizionate $c_{7,5}$, $c_{5,1}$, $c_{6,5}$, $c_{3,1}$, $c_{2,1}$, e $c_{4,3}$, in T_2 , 5 copulae condizionate $c_{7,1|5}$, $c_{5,3|1}$, $c_{6,1|5}$, $c_{4,1|3}$, e $c_{3,2|1}$ e, proseguendo sequenzialmente fino ad arrivare in T_6 , l'ultima copula bivariata condizionata $c_{6,4|72315}$.

Assumendo che le marginali siano distribuite secondo una v.c. Uniforme (0,1) la funzione di densità multivariata della copula può essere scritta:

$$c(y_1, \dots, y_n) = \prod_{i=1}^{n-1} \prod_{e \in E_i} c_{j(e), k(e)|D(e)}(F(y_{j(e)}|\mathbf{y}_{D(e)}), F(y_{k(e)}|\mathbf{y}_{D(e)})), \quad (2.1.4.1)$$

dove D rappresenta l'insieme degli elementi condizionanti e $\mathbf{y}_{D(e)}$ è il vettore $\mathbf{y} = (y_1, \dots, y_d)$ degli indici contenuti all'interno di $D(e)$.

Le funzioni di distribuzione condizionate che entrano come argomenti nell'equazione (2.1.4.1) vengono ricavate come segue (Joe, 1996):

per $j \in D$ e $D_{-j} := D \setminus j$

$$F(y_j|\mathbf{y}_D) = \frac{\partial C_{j,k|D-j}(F(y_j|\mathbf{y}_{D-j}), F(y_k|\mathbf{y}_{D-j}))}{\partial(F(y_k|\mathbf{y}_{D-j}))} \quad (2.1.4.2)$$

dove se $D = \{j\}$ la precedente equazione diventa:

$$F(y_j|y_k) = \frac{\partial C_{j,k}(F(y_j), (F y_k))}{\partial F(y_k)}$$

Nel caso di marginali distribuite in modo uniforme $u_j = F_j(y_j)$ l'equazione può essere riscritta come

$$h(u_j|u_k, \boldsymbol{\theta}_{jk}) = \frac{\partial C_{j,k}(u_j, u_k; \boldsymbol{\theta}_{jk})}{\partial u_k} \quad (2.1.4.3)$$

dove $\boldsymbol{\theta}_{jk}$ rappresenta la stima del parametro della copula bivariata applicata

alla coppia di variabili j, k .

La copula multivariata relativa alla sequenza T_1, \dots, T_{n-1} è costruita associando a ciascuna *edge* $e = j(e), k(e) | D(e)$ contenuta in E_i la copula bivariata $c_{j(e), k(e) | D(e)}$. La normalità asintotica degli Standard Error dei parametri stimati viene dimostrata da Haff (2013). Tuttavia, le covarianze asintotiche risultano molto complesse da calcolare e per tale ragione si ricorre alla metodologia bootstrap per stimare gli errori standard.

Dalla Definizione 2.1 si evince che una copula R-vine costruita su n variabili è un insieme nidificato di $n - 1$ *tree* tale per cui le *edges* del tree j diventano nodi del tree $j + 1$ (Figura 2.1.2). La condizione di prossimità assicura che due nodi nel tree $j + 1$ siano collegati da solo una *edge* se, e solo se, condividono un nodo comune nel tree precedente (j).

Come detto in precedenza, le sotto-rappresentazioni delle *Regular vine* sono le *Canonical vine* (C-vine) e le *Drawable-vine* (D-vine).

C-vine:

$$c(F_1(y_1), \dots, F_n(y_n)) = \prod_{i=1}^{n-1} \prod_{j=1}^{n-i} c_{i, i+j | 1, \dots, i-1}(F(y_i | y_1, \dots, y_{i-1}), F(y_{i+j} | y_1, \dots, y_{i-1}))$$

D-vine:

$$c(F_1(y_1), \dots, F_n(y_n)) = \prod_{i=1}^{n-1} \prod_{j=1}^{n-i} c_{j, j+i | (j+1): (j+i-1)}(F(y_j | y_{j+1}, \dots, y_{j+i-1}), F(y_{j+1} | y_{j+1}, \dots, y_{j+i-1}))$$

La rappresentazione grafica di queste due copulae è ben diversa tra loro. La prima è costituita da un nodo centrale al quale vengono collegati tutti gli altri nodi (grafico a stella), mentre la seconda segue un percorso sequenziale lineare. Un esempio grafico viene riportato in Figura 2.1.3 di seguito. Per una trattazione più dettagliata di tali modelli si veda Joe (1997).

Definizione 2.2 *Data la matrice triangolare bassa $M = (m_{j,k})_{j,k=1,\dots,n}$, il j -esimo set di vincoli per M è dato da*

$$C_M(i) = \{(\{m_{j,j}, m_{i,j}\}, D) | i = j + 1, \dots, n, D = \{m_{i+1,j}, \dots, m_{n,j}\}\} \quad (2.1.4.5)$$

per $j = 1, \dots, n - 1$. Se $i = n$ allora $D = \emptyset$. L'insieme di tutti i vincoli della matrice M corrisponde all'unione $CM = C_M(1) \cup \dots \cup C_M(n - 1)$. Dati gli elementi $(\{m_{j,j}, m_{i,j}\}, D) \in CM$ è possibile definire $\{m_{j,j}, m_{i,j}\}$ come set condizionato e D come set condizionante.

La matrice 2.1.4.4 costituisce la rappresentazione matriciale della Figura 2.1.2. Per dimostrare l'idea prendiamo come esempio l'elemento sulla diagonale della prima colonna $m_{1,1} = 6$ e l'elemento della colonna corrispondente $m_{3,1} = 7$. In accordo con la Definizione 2.2 abbiamo l'insieme $(\{6, 7\}, \{2, 3, 1, 5\}) \in CM$ dove la coppia $\{6, 7\}$ costituisce l'insieme vincolato, mentre $\{2, 3, 1, 5\}$ rappresenta l'insieme vincolante. La chiave di lettura della matrice R-vine segue uno schema particolare: dal basso verso l'alto e da destra verso sinistra. Infatti le *edge* presenti nel primo tree sono costituite dalla combinazione degli elementi presenti sulla diagonale ed i corrispondenti elementi in colonna dell'ultima riga. Per T_2 le edge corrispondono alle coppie $\{6, 1\}, \{7, 1\}, \{5, 3\}, \{4, 1\}, \{3, 2\}$ dove il set condizionante è rappresentato dagli elementi presenti nella riga n partendo dalla colonna 1 fino alla colonna $n - 1$. Continuando con lo stesso metodo di lettura si riesce facilmente a costruire la struttura ad albero della Figura 2.1.2.

Morales-Nápoles (2008) individuano un numero di possibili diverse rappresentazioni R-vine in n dimensioni pari a $\binom{n}{2} \times (n - 2)! \times 2^{\binom{n-2}{2}}$. Occorre pertanto un metodo di selezione delle singole coppie di variabili.

L'idea di Berg and Ass (2009) e Dißmann, Brechmann, Czado, and Kurowicka (2013) è quella di catturare le relazioni di dipendenza più forti nei primi *tree* attraverso l'utilizzo del metodo *maximum spanning tree*. A tal proposi-

to viene calcolato, per ciascuna coppia di variabili, un indice di dipendenza¹ (generalmente l'indice di Kendall)

$$\max_{\text{edges } e = (i,j) \text{ in spanning tree}} \sum |\delta_{i,k}|, \quad \text{con } i \neq j$$

dove $\delta_{i,k}$ rappresenta l'indice di Kendall e con il termine *spanning tree* ci si riferisce alla sequenza di variabili che rispettano la condizione di prossimità indicata nella definizione 2.1.4.

viene utilizzato il valore assoluto dell'indice di dipendenza di modo da dare lo stesso peso a valori positivi e negativi.

2.1.5 Stima modello R-vine a più stadi

Nel precedente paragrafo è stata descritta una procedura di costruzione di un modello con copula R-vine. La stima di tale modello può essere compiuta attraverso una procedura di Massima Verosimiglianza a più stadi.

È noto in letteratura che lo stimatore di Massima verosimiglianza $\hat{\boldsymbol{\theta}}_n$ del vettore k -dimensionale dei parametri $\boldsymbol{\theta}$, ottenuto da un campione di n osservazioni e sotto determinate condizioni di regolarità (Bickel and Doksum, 2007), è consistente e asintoticamente normale:

$$\sqrt{n} I(\boldsymbol{\theta})^{1/2} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(0, \mathbf{I}d_k), \quad \text{con } n \rightarrow \infty$$

dove $\boldsymbol{\theta} \in \mathbb{R}^k$ rappresenta il vettore dei veri parametri, $\mathbf{I}d_k$ la matrice identità di dimensione $k \times k$ e $\mathcal{I}(\boldsymbol{\theta})$ la matrice di informazione di Fisher (attesa).

$$\mathcal{I}(\boldsymbol{\theta}) = -\mathbb{E}_{\boldsymbol{\theta}} \left[\left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} l(\boldsymbol{\theta} | \mathbf{X}) \right)_{i,j=1,\dots,k} \right] = \mathbb{E}_{\boldsymbol{\theta}} \left[\left(\frac{\partial}{\partial \theta_i} l(\boldsymbol{\theta} | \mathbf{X}) \frac{\partial}{\partial \theta_j} l(\boldsymbol{\theta} | \mathbf{X}) \right)_{i,j=1,\dots,k} \right] \quad (2.1.5.1)$$

¹Indice di correlazione di rango. Date due variabili casuali esso viene definito come

$$\frac{C - D}{C + D}$$

dove C rappresenta il numero di coppie concordanti e D il numero di coppie discordanti.

con $l(\boldsymbol{\theta}|\mathbf{x})$ che rappresenta la funzione log-verosimiglianza del parametro $\boldsymbol{\theta}$ valutata in $\mathbf{X} = \mathbf{x}$.

In un campione finito di n osservazioni indipendenti (x_1, \dots, x_n) Efron and Hinkley (1978) dimostrano che la matrice di informazione di Fisher (2.1.5.1) può essere sostituita prendendo la semplice media campionaria ed ottenendo così la matrice di informazione osservata:

$$\mathcal{I}_n(\hat{\boldsymbol{\theta}}_n) = \left[\left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \sum_{i=1}^n l(\boldsymbol{\theta}|x_i) \right)_{i,j=1,\dots,k} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_n}$$

Seguendo l'impostazione di Hobæk Haff (2013), la matrice di covarianza asintotica del vettore dei parametri $\boldsymbol{\theta}$ può essere scomposta in:

$$\mathcal{V}_\theta = \mathcal{V}_{\text{dependence}} + \mathcal{V}_{\text{margins}} \quad (2.1.5.2)$$

dove $\mathcal{V}_{\text{dependence}} = \mathcal{I}(\boldsymbol{\theta})^{-1}$ della stima di Massima Verosimiglianza e la seconda parte è pari a zero solo se sono note le distribuzioni marginali. Utilizzando le definizioni introdotte nel Paragrafo 2.1.4, la funzione di log-verosimiglianza per la stima dei parametri dell' i -esimo tree è data da

$$l_i(\boldsymbol{\theta}|\mathbf{x}) = \sum_{e \in E_i} \log(c_{j(e),k(e)|D(e)}(F_{j(e)|D(e)}(x_{j(e)}|\mathbf{x}_{D(e)}), F_{k(e)|D(e)}(x_{k(e)}|\mathbf{x}_{D(e)})|\boldsymbol{\theta}_e))$$

dove il vettore dei parametri $\boldsymbol{\theta}_i$ è $\{\theta_e | e \in E_i\}$. Il parametro θ_1 viene stimato massimizzando la funzione $l_1(\boldsymbol{\theta}_1|\mathbf{x})$ e la sua stima $\hat{\boldsymbol{\theta}}_1$ viene utilizzata all'interno della 2.1.4.2 per costruire le pseudo osservazioni che sono gli argomenti della funzione $l_2(\boldsymbol{\theta}_2|\mathbf{x})$. Si procede con la massimizzazione di quest'ultima al fine di ottenere lo stimatore $\boldsymbol{\theta}_2$. Tale procedura viene eseguita fino a che tutti i parametri della copula R-vine non sono stati stimati.

La matrice di covarianza asintotica delle stime ottenute è data da:

$$\mathcal{V}_{\text{dependence}} = \mathcal{J}_\theta^{-1} \mathcal{K}_\theta (\mathcal{J}_\theta^{-1})^T \quad (2.1.5.3)$$

con

$$\mathcal{K}_\theta = \begin{pmatrix} \mathcal{K}_{\theta,1,1} & & & & \\ \vdots & \ddots & & & \\ 0^T & \dots & \mathcal{K}_{\theta,n-2,n-2} & & \\ 0^T & \dots & 0^T & & \mathcal{K}_{\theta,n-1,n-1} \end{pmatrix}, \quad (2.1.5.4)$$

$$\mathcal{J}_\theta = \begin{pmatrix} \mathcal{J}_{\theta,1,1} & & & & \\ \vdots & \ddots & & & \\ \mathcal{J}_{\theta,n-2,1} & \dots & \mathcal{J}_{\theta,n-2,n-2} & & \\ \mathcal{J}_{\theta,n-1,1} & \dots & \mathcal{J}_{\theta,n-1,n-2} & \mathcal{J}_{\theta,n-1,n-1} \end{pmatrix} \quad (2.1.5.5)$$

dove $\mathcal{K}_{\theta,i,j} = \mathbb{E} \left[\left(\frac{\partial_i(\theta_i, \dots, \theta_1 | \mathbf{X})}{\partial \theta_i} \right) \left(\frac{\partial_j(\theta_j, \dots, \theta_1 | \mathbf{X})}{\partial \theta_j} \right)^T \right]$ e $\mathcal{J}_{\theta,i,j} = -\mathbb{E} \left[\frac{\partial^2 l_i(\theta_i, \dots, \theta_1 | \mathbf{X})}{\partial \theta_i \partial \theta_j} \right]$, $i, j = 1, \dots, n-1$.

Nel caso di distribuzioni marginali non note l'equazione 2.1.5.2 può essere scritta come:

$$\mathcal{V}_\theta^{SPP} = \mathcal{J}_\theta^{-1} \mathcal{K}_\theta (\mathcal{J}_\theta^{-1})^T + \mathcal{J}_\theta^{-1} \mathcal{B}_\theta^{SPP} (\mathcal{J}_\theta^{-1})^T \quad (2.1.5.6)$$

dove la matrice \mathcal{B}_θ^{SPP} quantifica l'effetto dell'utilizzo delle funzioni di distribuzioni empiriche in sostituzione delle distribuzioni marginali reali.

Sotto le condizioni (A.2) - (A.5) di Tsukahara (2005), Hobæk Haff (2013) dimostra che lo stimatore semiparametrico $\hat{\theta}^{SPP}$ è consistente e asintoticamente normale:

$$\sqrt{n}(\hat{\theta}^{SPP} - \theta) \xrightarrow{d} N(0, \mathcal{V}_\theta^{SPP}) \quad (2.1.5.7)$$

Schepsmeier and Stöber (2014) ricavano nel loro lavoro le espressioni per la matrice di informazione attesa ed osservata per diverse famiglie di copulae bivariate fornendo anche le derivate della loro funzione di verosimiglianza. Data la procedura gerarchica appena descritta tale lavoro risulta essere di

particolare importanza nell'analisi inferenziale delle *Vine Copulae* e nel calcolo degli errori standard dei parametri stimati.

Goodness-of-Fit (GoF) test

Testare la bontà di adattamento dei modelli (*Goodness-of-Fit* (GoF)) con copula risulta essere di grande rilevanza in quanto tali modelli potrebbero presentare dei problemi di errata specificazione.

I due test maggiormente utilizzati sono i test Kolmogorov-Smirnov (KS) e Cramér-von-Mises (CvM). Entrambe si basano sulla comparazione tra la CDF (Cumulative Distribution Function) della copula generata dal modello con la copula empirica di seguito definita:

$$\hat{C}_T(u) = \frac{1}{T} \sum_{t=1}^T \prod_{i=1}^n \mathbf{1}(\hat{U}_{it} \leq u_{it})$$

dove $\hat{U}_{it} = \hat{F}_i(\hat{\epsilon}_{it})$ con \hat{F}_i che rappresenta la funzione di distribuzione empirica (*Empirical Distribution Function* (EDF)) e $\hat{\epsilon}_{it}$ sono i residui standardizzati stimati. Genest and D. (2009) individuano una versione più potente del test CvM basata su un bootstrap parametrico in cui il p-value viene approssimato attraverso il semplice calcolo della media bootstrap Genest and B. (2008). Tale procedura è però scarsamente utilizzata in quanto richiede un elevato numero di simulazioni comportando uno notevole sforzo computazionale. Al fine di ovviare a questo problema Kojadinovic and Yan (2011) introducono il *Multiplier GoF* test basato sul *Multiplier Central Limiti Theorems* (van der Vaart and Wellner, 1996) e sulla copula empirica. L'idea è quella di comparare la copula empirica $\hat{C}_T(u)$ con lo stimatore parametrico $C_{\hat{\theta}}(u)$ ottenuto sotto l'ipotesi nulla $H_0 : C \in \mathcal{C}_0 = \{C_{\theta} : \theta \in \mathcal{O}\}$, dove \mathcal{O} è un sottoinsieme di \mathbb{R}^q per qualunque $q = \{1, 2, 3, \dots\}$. Fang, Madsen, and Liu (2014) dimostrano come tale procedura risulta essere molto più veloce dei metodi basati su bootstrap parametrici.

I test appena citati si basano sull'assunzione che la Copula condizionata (2.1.0.4) sia costante nel tempo e quindi risultano essere inadeguati su modelli con copula dinamici. La soluzione a tale problematica è data dall'utilizzo della trasformazione di Rosenblatt, ovvero la versione multivariata della *probability integral transformation* (PIT) (Diebold, Hahn, and Tay, 1999), (Rémillard, 2010). In questo caso i dati vengono inizialmente trasformati, divenendo così indipendenti e distribuiti secondo una variabile casuale Uniforme (0,1) e successivamente vengono applicati i test KS e CvM sui dati trasformati.

I test appena descritti possono essere utilizzati, accanto ai criteri informativi AIC e BIC, nella selezione delle diverse copulae bivariate all'interno della procedura R-vine. Fang, Madsen, and Liu (2014) dimostrano come il criterio informativo di Akaike risulta essere più efficiente ed accurato rispetto al Multiplier GoF.

2.1.6 Vantaggi e Svantaggi

Esistono diversi vantaggi, soprattutto in campo finanziario, dovuti all'utilizzo di modelli con copula. Uno dei principali è che tali modelli riescono facilmente a modellare l'eventuale presenza di dinamiche di dipendenza non lineari esistenti all'interno di portafogli finanziari. Inoltre la Copula permette di collegare tra loro diverse variabili che potenzialmente seguono delle distribuzioni marginali differenti e di isolare la struttura di dipendenza esistente tra le stesse. Ciò rende tale strumento il più flessibile in letteratura, non dovendo effettuare alcuna assunzione né sulle distribuzioni marginali né sulla distribuzione congiunta. In questo modo la struttura di dipendenza risulta essere totalmente slegata ed indipendente rispetto alla scelta delle distribuzioni marginali. Questa caratteristica è molto utile in applicazioni riguardanti problemi di asset allocation e risk hedging in quanto le singole serie dei rendimenti potrebbero possedere distribuzioni diverse ed avere un andamento congiunto non standard.

Un'altra caratteristica fondamentale è che tale strumento è, a differenza della semplice correlazione lineare, invariante rispetto a trasformazioni sui dati. Ciò significa che la copula riesce ad estrarre il comportamento congiunto tra le varie variabili indipendentemente dalla scala utilizzata (ad esempio i logaritmi dei rendimenti o altre trasformazioni).

In ambito multivariato, assumere una predeterminata funzione di distribuzione delle struttura di dipendenza può risultare estremamente semplificatrice. Attraverso l'utilizzo della metodologia PCC (Paragrafo 2.1.3), per la costruzione di copule multivariate, questo limite viene superato rendendo possibile la costruzione una qualsiasi copula multivariata attraverso una semplice aggregazione di copule bivariate. Ciò permette di ottenere una funzione di dipendenza più flessibile rispetto ai modelli già esistenti in letteratura.

Gli svantaggi principali generati dall'utilizzo di tale strumento si trovano principalmente in ambito multivariato. Come descritto in dettaglio nella sezione 2.1.2, la stima *Full Maximum Likelihood* di un modello con copula comporta delle notevoli difficoltà computazionali dovute alla stima simultanea dei parametri delle marginali e della Copula. Il modello PCC, approssima la funzione di verosimiglianza attraverso una fattorizzazione sequenziale e, per tale ragione, comporta una perdita di efficienza nella stima, anche se minima. L'utilizzo di tale metodologia genera inoltre un incremento notevole del numero dei parametri (soprattutto per un numero elevato di variabili) comportando conseguentemente ad un aumento nei tempi di stima.

Questa problematica può essere risolta facendo ricorso alla computazione parallela. Dopo aver definito la sequenza di variabili in ogni tree, la procedura di stima mette a confronto diverse tipologie di copule bivariate per poi scegliere la migliore sulla base di criteri prestabiliti. Avendo a disposizione un numero di processori maggiori di due è possibile eseguire stime simultanee sia per le diverse coppie di variabili e/o diverse famiglie di copule. In questo modo i tempi macchina vengono ridotti drasticamente.

Un ulteriore svantaggio riguarda l'arbitrarietà nei criteri che definiscono

la struttura dell'albero. Infatti nella costruzione delle Vine Copulae viene spesso utilizzato il criterio del *minimum spanning tree* dove viene massimizzata la somma dei valori assoluti degli indici di dipendenza di Kendall tra coppie di variabili. La scelta dell'indice di Kendall è del tutto arbitraria in quanto potrebbero essere utilizzati indici diversi di come ad esempio l'indice di correlazione di Pearson, che condurrebbero ad una struttura ad albero diversa.

2.2 Modelli di portafoglio con Copula

2.2.1 Patton (2004)

Nel suo lavoro Patton prende in esame un portafoglio con due indici azionari ed un tasso privo di rischio: indice azionario di società a larga capitalizzazione, indice azionario di società a bassa capitalizzazione ed il Treasury Bill. La somma dei pesi può essere inferiore ad uno, questo significa che l'investitore può destinare una parte del capitale in qualche attività priva di rischio. Patton utilizza un'equazione per la media condizionata, come funzione dei 12 ritardi dei rendimenti, del tasso risk-free, del default spread e del dividend yield. Inoltre viene utilizzato un Threshold-ARCH(1,1) (Zakoian, 1994) per la varianza con le stesse tre variabili utilizzate per la media, ritardate di un periodo.

Per catturare sia l'asimmetria sia la curtosi c'è bisogno di modellare le distribuzioni marginali dei residui e della copula. Patton utilizza la distribuzione skewed-t come suggerita da Hansen (1994) ed analizzata da Jondeau and Rockinger (2003). I parametri dell'asimmetria ed i gradi di libertà vengono ipotizzati come funzione del tasso privo di rischio ritardato, del default spread ritardato, del dividend yield ritardato e delle previsioni del modello riguardanti la media e la varianza. Seguendo l'impostazione di Hansen, questi parametri vengono trasformati per poter far rientrare l'asimmetria tra -1 e 1 ed i gradi di libertà superiori a 2.

Patton prende in considerazione diverse copulae, ma trova che la Gumbel risulta essere la migliore in quanto con essa ottiene i valori migliori nella log-verosimiglianza e nei criteri informativi.

Egli, nel suo lavoro, ha effettuato diverse prove di allocazione del capitale:

- concentrazione in un unico asset;
- pesi distribuiti egualmente tra le attività;
- pesi calcolati utilizzando distribuzioni dei rendimenti non condizionate:

$$w_{t+1} = \arg \max_w \frac{\sum_{j=1}^t U(1 + w_x X_j + w_y Y_j)}{t} \quad (2.2.1.1)$$

dove w rappresenta i pesi e t indica il tempo.

- calcolo dei pesi utilizzando le specificazioni di cui sopra, con ipotesi di normalità dei residui;
- calcolo dei pesi con marginali condizionate distribuite come una skewed-t ed utilizzando la copula gaussiana.

Patton, nel calcolo dei pesi ottimali, ipotizza una funzione di utilità CRRA (Constant Relative Risk Aversion). Quest'ultima viene spesso utilizzata in applicazioni di ambito finanziario e può essere pensata come una stima prudentiale dei guadagni ottenibili evitando l'utilizzo della distribuzione normale. L'autore utilizza 5 diversi gradi di avversità al rischio $\gamma : 1, 3, 7, 10, 20$. Il lavoro prende in considerazione rendimenti mensili dal 1954 al 1999 lasciando il periodo dal 1990 al 1999 per un'analisi out-of-sample. I risultati dello studio mostrano che il miglior modello è quello che utilizza marginali distribuite come skewed-t, la copula Gumbel ruotata e con vincoli sulle vendite allo scoperto. Questo modello fornisce i risultati migliori quando si è in presenza di un'elevata avversione al rischio. Inoltre risulta essere migliore rispetto al modello media-varianza anche in assenza di vincoli per le vendite allo scoperto perché in tal caso è di maggiore importanza il riuscire a catturare l'asimmetria sia

nella distribuzione sia nella dipendenza tra le due attività. I pesi calcolati con tale modello risultano essere meno variabili in relazione alle variazioni nei rendimenti attesi, ma molto sensibili a variazioni di dipendenza tra le attività.

2.2.2 Riccetti (2013)

Nel lavoro “A copula–GARCH model for macro asset allocation of a portfolio with commodities” l’autore si riallaccia ad un suo stesso lavoro del 2010 ed al lavoro di Patton (2004). Riccetti (2010) conclude che, l’utilizzo della copula all’interno di modelli di allocazione di portafoglio, risulta essere utile nella scelta delle macro-asset-class (stock e bond) specialmente quando l’investitore è particolarmente avverso al rischio. Invece, non è di alcuna utilità nella scelta dei pesi in un portafoglio composto da soli titoli azionari. La copula sembra essere di una qualche utilità nei casi in cui all’interno di un portafoglio vi sia la presenza di un’attività finanziaria che possiede una scarsa correlazione con le altre attività.

In questo lavoro, l’autore, vuole indagare se tali risultati vengono confermati anche in presenza di un indice sulle commodities. Egli utilizza un modello con copula per l’allocazione tra due, tre e quattro asset analizzando varie combinazioni per testare se l’utilizzo delle copulae riesce ad incrementare l’utilità dell’investitore.

Le assunzioni del modello sono:

- momento primo non condizionato;
- varianza modellata con un GARCH(1,1) con termini di innovazione estratti dalla distribuzione *Skewed-t*;
- il comportamento congiunto dei residui viene modellato con la copula che ottiene il valore di log-verosimiglianza più elevato.

Per semplificare il lavoro vengono esclusi costi di transazione, costi di riallocazione del capitale e commissioni per le vendite allo scoperto.

A differenza di Patton, che utilizza rendimenti mensili, in questo caso vengono presi in considerazione i rendimenti giornalieri con lo scopo di ottenere allocazioni efficienti di capitale su base settimanale, mensile ed annuale. Il periodo va dal 1 Gennaio 1996 al 31 Agosto 2010 con gli ultimi 3 anni utilizzati per l'analisi out-of-sample. Nell'analisi bivariata della struttura di dipendenza vengono utilizzate diverse copulae: Clayton, Gumbel, Frank, Gaussiana e t-Student, mentre nel multivariato viene presa in considerazione la copula *Canonical Vine*.

Nella creazione del portafoglio vengono seguite alcune fasi:

- stima dei parametri con la Massima-Verosimiglianza;
- scelta dei pesi ottimali al fine di massimizzare l'utilità dell'investitore utilizzando la funzione di utilità CRRA (come in Patton (2004)):

$$U(\gamma) = \begin{cases} (1 - \gamma)^{-1} [P_0(1 + R)]^{1-\gamma} & \text{se } \gamma \neq 1 \\ \log[P_0(1 + R)] & \text{se } \gamma = 1 \end{cases}$$

con R_{port} come fattore di capitalizzazione del portafoglio e $P_0 = 1$ come investimento iniziale. I valori di γ sono: 2, 5, 10 e 15.

- si effettuano 5000 prove di allocazione per diversi orizzonti temporali (5 giorni per un'allocazione settimanale).

Da tutta questa analisi emergono dei risultati contrastanti: da una parte viene confermata l'utilità dell'utilizzo delle copulae nella scelta delle macro-asset-class come in Riccetti (2010). Considerando due sole macro-asset-class, il modello con copula risulta essere il migliore per ciascun livello di avversione al rischio considerato. Anche nel caso di 3 attività finanziarie il portafoglio costruito con copula ottiene performance migliori rispetto al portafoglio media-varianza. D'altro canto, però, analizzando 4 attività finanziarie scarsamente correlate tra loro, il modello con copula non risulta essere il migliore in assoluto. Dalla comparazione del modello con copula con un modello naive, nella

porzione out-of-sample del campione, i due modelli ottengono più o meno gli stessi risultati. Il modello equipesato si basa sull'assunzione di indipendenza dei residui ed è decisamente più semplice e più veloce da applicare, e ciò indica una scarsa utilità del modello con copula quando vengono considerate tutte attività finanziarie scarsamente correlate tra loro.

In conclusione, il modello con copula ottiene le migliori performance out-of-sample quando all'interno del portafoglio vengono considerate:

- due attività finanziarie non correlate tra loro;
- due o più attività finanziarie correlate ed una non correlata, come mostrato anche da Riccetti (2010) e Patton (2004).

L'autore, al fine di ottenere previsioni out-of-sample migliori, ipotizza per degli studi futuri l'utilizzo di modelli più complessi che prendano in considerazione ad esempio delle copulae con parametri dinamici.

Tale considerazione ha fornito uno spunto per lo studio svolto in questo lavoro.

2.2.3 Truncated and simplified regular vines (Brechmann, Czado, and Aas, 2012)

La flessibilità che contraddistingue i modelli con copula R-vine porta inevitabilmente ad un incremento della loro complessità soprattutto quando viene preso in esame un numero elevato di variabili. Questa problematica provoca un aumento dei tempi-macchina per la stima dell'intero modello ed un notevole consumo di CPU. Tale modello cerca di ridurre questo problema indagando la possibilità di un troncamento o semplificazione della stima sequenziale ad un determinato livello (*tree*).

La copula R-vine, come descritto in precedenza, è composta da una serie di copulae bivariate. La tecnica di semplificazione ad un determinato livello K , dell'intera struttura R-vine, consiste nella sostituzione di tutte le copulae bivariate presenti all'interno del (*tree*) situato ad un livello maggiore o uguale

a K con copulae Gaussianne. La scelta della copula Gaussiana è dettata dalla rapidità del processo di stima e dalla facile interpretazione del parametro al suo interno.

La tecnica di troncamento funziona in modo analogo, con la differenza che, in questo caso, viene utilizzata la *independence copula* come copula sostitutiva.

Indichiamo con $tRV(K)$ la copula R-vine troncata a livello K e con $sRV(K)$ la copula semplificata. Le funzioni di densità di queste due copulae sono:

$$c_{tRV(K)}(y_1, \dots, y_n | \boldsymbol{\theta}_{tRV(K)}) = \prod_{i=1}^K \prod_{e \in E_i} c_{j(e),k(e)|D(e)}(F(y_{j(e)} | \mathbf{y}_{D(e)}), F(y_{k(e)} | \mathbf{y}_{D(e)})) \quad (2.2.3.1)$$

con $\boldsymbol{\theta}_{tRV(K)} = \{\boldsymbol{\theta}_{j(e),k(e)|D(e)} : e \in E_i, i = 1, \dots, K\}$, dove $\boldsymbol{\theta}_{j(e),k(e)|D(e)}$ indica il parametro/i della copula $c_{j(e),k(e)|D(e)}$.

$$c_{sRV(K)}(y_1, \dots, y_n | \boldsymbol{\theta}_{sRV(K)}) = \left[\prod_{i=1}^K \prod_{e \in E_i} c_{j(e),k(e)|D(e)} \right] \times \left[\prod_{i=K+1}^{n-1} \prod_{e \in E_i} c_{j(e),k(e)|D(e)}^\rho \right], \quad (2.2.3.2)$$

dove $c_{j(e),k(e)|D(e)}^\rho$ rappresenta la funzione di densità della copula bivariata Gaussiana con parametro $\rho_{j(e),k(e)|D(e)}$. Il parametro $\boldsymbol{\theta}_{sRV(K)}$ indica l'insieme dei parametri di $c_{sRV(K)}$ ovvero

$$\begin{aligned} \boldsymbol{\theta}_{sRV(K)} = & \{\boldsymbol{\theta}_{j(e),k(e)|D(e)} : e \in E_i, i = 1, \dots, K\} \\ & \cup \{\rho_{j(e),k(e)|D(e)} : e \in E_i, i = K + 1, \dots, n - 1\}, \end{aligned}$$

La selezione del livello K viene effettuata partendo dall'assunzione che la copula $tRV(K)$ è contenuta all'interno della copula $tRV(K + 1)$. Viene effettuata una stima sequenziale partendo da un livello $K = 1$ ed aggiungendo un *tree* alla volta valutando se vi è un incremento significativo di adattamento del modello ai dati. A tal proposito vengono confrontati i criteri informativi AIC/BIC dei due modelli e viene effettuato il test di rapporto di verosimiglianza proposto da Vuong (1989). Il procedimento è sostanzialmente lo stesso per la copula

$sRV(K)$ ma in questo caso, contrariamente a quanto visto sopra, i due modelli non sono nidificati. Il confronto viene comunque effettuato attraverso i criteri informativi anche se il loro utilizzo, in questa situazione, può comportare qualche problema (Ripley 2008, pp. 34-35).

L'intero studio viene sviluppato su un dataset di 19 variabili di mercato Norvegesi ed internazionali. I risultati mostrano che la maggior parte della dipendenza viene catturata all'interno dei primi 6 tree e ciò significa che la copula R-vine può essere troncata ad un livello $K = 6$. La procedura di semplificazione invece indica un livello pari a 2 ($K = 2$), quindi le copulae bivariate presenti nei tree successivi possono essere sostituite dalla più generale copula Gaussiana senza una rilevante perdita di efficienza.

Questa procedura aumenta la flessibilità ed il campo di utilizzo dei modelli con copula R-vine riducendo notevolmente i tempi di stima e rendendoli così maggiormente utilizzabili su database di grandi dimensioni.

2.2.4 Dynamic D-vine model (Almeida, Czado, and Manner, 2016)

Questo lavoro è stato uno dei primi ad utilizzare le copulae al fine di analizzare l'eventuale presenza di una struttura di dipendenza variabile nel tempo attraverso su un numero elevato di variabili.

Gli autori, nel loro studio, sviluppano due modelli diversi combinando il metodo del pair-copula construction nel primo caso con una copula autoregressiva stocastica (SCAR, stochastic autoregressive copula), mentre nel secondo con un modello GAS (Generalized Autoregressive Score).

Per semplicità, prendiamo in considerazione due serie storiche $(y_{i,t}, y_{j,t})$ con $t = 1, \dots, T$ la cui funzione di distribuzione è data da:

$$(y_{i,t}, y_{j,t}) \sim C(u_{jt}, u_{jt}; \theta_t^{ij})$$

con $u_{it}, u_{jt} \in [0, 1]$ che rappresentano i valori trasformati delle variabili y_{it} ed y_{jt} attraverso l'utilizzo della loro funzione di distribuzione inversa e $\theta_t^{i,j} \in \Theta$ indica il parametro time-varying della copula C.

Nell'approccio SCAR (Hafner and Manner (2012) e Almeida and Czado (2012)) il parametro della copula segue un processo autoregressivo di ordine (1) dato da

$$\theta_t^{ij} = \mu_{ij} + \phi_{ij}(\theta_{t-1}^{ij} - \mu_{ij}) + \sigma_{ij}z_{ij,t},$$

dove $z_{ij,t}$ rappresenta i termini di innovazioni distribuiti secondo una normale standardizzata. Ulteriori assunti sono $|\phi_{ij}| < 1$ e $\sigma_{ij} > 0$ indispensabili ai fini della stazionarietà e dell'identificazione del modello.

La specificazione della dinamicità studiata attraverso l'utilizzo di un modello GAS (Creal, Koopman, and Lucas (2013)) presenta la seguente formulazione:

$$\theta_t^{ij} = \omega_{ij} + \phi_{ij}\theta_{t-1}^{ij} + \delta_{ij}s_t^{ij}, \quad (2.2.4.1)$$

dove s_t^{ij} costituisce il vettore dello score al quale viene applicato un fattore di scala

$$s_t^{ij} = S_{ij,t} \nabla_{ij,t},$$

in cui la funzione score è pari a

$$\nabla_{ij,t} = \frac{\partial \ln c(y_{i,t}, y_{j,t}; \boldsymbol{\omega}_{ij}, F_{t-1})}{\partial \theta_t^{ij}}$$

dove $\boldsymbol{\omega}_{ij} = (\omega_{ij}, \phi_{ij}, \delta_{ij})$. $S_{ij,t}$ indica una matrice di scala ed è data dalla radice quadrata dell'inversa della matrice di informazione.

Combinando la formulazione D-vine con i modelli dinamici appena descritti, la funzione di densità della copula multivariata può essere riscritta come

$$c(y_1, \dots, y_n; \boldsymbol{\theta}_t) := \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{l(i,j)}(F(y_i | \mathbf{y}_{i+1:i+j-1}; \boldsymbol{\theta}_t^{l(i,j)}), F(y_{i+j} | \mathbf{y}_{i+1:i+j-1}); \boldsymbol{\theta}_t^{l(i,j)}), \quad (2.2.4.2)$$

dove $l(i, j) = i, i+j | i+1 : i+j-1$ e $\boldsymbol{\theta}_t = \{\boldsymbol{\theta}_t^{l(i,j)}; j = 1, \dots, n-1, i = 1, \dots, n-1\}$ è il vettore dei parametri time-varying della copula. La tipologia di copula corrispondente a $l(i, j)$ può essere scelta in maniera del tutto indipendente da altri indici $l(r, s)$.

La stima del modello D-vine-GAS viene effettuata con il metodo di massima verosimiglianza attraverso la procedura a due stadi descritta in precedenza. In questo caso i parametri delle copulae bivariate vengono ricavati utilizzando la formulazione (2.2.4.1) per poi essere inseriti nell'equazione (2.1.4.3) per il calcolo delle pseudo osservazioni da utilizzare nei tree successivi.

La stima del modello SCAR-DVine risulta essere più difficoltosa rispetto alla precedente. L'interesse principale si concentra sulla stima del vettore dei parametri $\boldsymbol{\omega} := (\mu, \phi, \sigma)$. Poniamo $\Theta = \{\boldsymbol{\theta}_t\}_{t=1}^T$. La funzione di verosimiglianza del vettore dei parametri $\boldsymbol{\omega}$ può essere ottenuta integrando la seguente funzione:

$$L(\mathbf{y}_i, \mathbf{y}_j; \boldsymbol{\omega}) = \int \prod_{t=1}^T f(y_{i,t}, y_{j,t}, \boldsymbol{\theta}_t | \boldsymbol{\theta}_{t-1}, \boldsymbol{\omega}) d\Theta$$

La risoluzione di questo integrale viene effettuata attraverso delle simulazioni di Monte Carlo utilizzando la tecnica *efficient importance sampling* introdotta da Richard and Zhang (2007).

Le performance dei due modelli vengono analizzate all'interno di due diversi dataset. Il primo è composto dalla serie dei rendimenti settimanali dell'indice MSCI per gli Stati Uniti, Regno Unito, Europa, Canada e Giappone per il periodo 7/Gen/1971 - 10/Ott/2013. Le tipologie di copulae utilizzate sono la Gaussiana, la Gumbel e la Clayton con le rispettive versioni ruotate e time-varying.

La scelta delle copulae all'interno dei primi tree è ricaduta ampiamente sulle versioni time-varying sia per il modello SCAR-DVine sia per il GAS-DVINE. La reale utilità dei due modelli viene esaminata attraverso il confronto con il modello DCC con copula Gaussiana ed un modello con copula R-vine (Tabella

1.4)

Tabella 2.2.1: In-and out-of-sample fit

Model	LL	BIC	Pred. LL	# param.
SCAR-DVine	2645.39	-5153.96	515.48	18
GAS-DVine	2912.41	-5093.15	464.43	26
DCC	2030.75	-3970.29	479.50	12
RVine (Constant)	2089.42	-4057.22	404.74	16

La tabella presenta la bontà di adattamento in-sample e out-of sample dei modelli analizzati in termini di log-verosimiglianza (LL), BIC e log-verosimiglianza prevista (Pred. LL). Le colonne LL e BIC coprono il periodo in-sample dal 07/Gen/1971 al 30/Giu/2009. Pred. LL viene calcolata per il periodo out-of-sample dal 6/Lug/2009 al 10/Ott/2013. L'ultima colonna indica il numero di parametri di ogni modello.

Dalla tabella 1.4 si nota subito che, nel porzione in-sample del campione, i due modelli analizzati raggiungono dei risultati decisamente migliori rispetto al modello DCC ed R-Vine. Nel porzione out-of-sample, invece, il DCC ottiene delle buone performance, ma il modello SCAR risulta ancora essere il migliore

Il secondo dataset è costituito dalle serie dei rendimenti giornalieri di 29 asset quotati nel DAX30 nel periodo dal 1/Gen/2008 al 31/Dic/2013. I primi 4 anni rappresentano il periodo in-sample mentre gli anni 2012 e 2013 sono utilizzati per valutare le previsioni.

A differenza di quanto visto precedentemente, in questo dataset il modello con copula RVine ottiene degli ottimi risultati raggiungendo il valore più elevato di log-verosimiglianza ma a causa dell'elevato numero di parametri al suo interno il criterio informativo BIC risulta essere più alto rispetto al modello SCAR-DVine. I due modelli rimanenti ottengono risultati decisamente peggiori.

Lo studio appena riportato mette in luce la necessità di modelli che prendono in considerazione una qualche dinamica temporale dei parametri di dipendenza.

Tabella 2.2.2: In-and out-of-sample fit

Model	LL	BIC	Pred. LL	# param.
SCAR-DVine	10195.17	-18986.86	5567.36	203
GAS-DVine	9200.46	-17294.72	4378.58	160
DCC	9984.34	-17147.47	4513.32	408
RVine (Constant)	10587.76	-18368.54	4128.80	406

La tabella presenta la bontà di adattamento in-sample e out-of sample dei modelli analizzati in termini di log-verosimiglianza (LL), BIC e log-verosimiglianza prevista (Pred. LL). Le colonne LL e BIC coprono il periodo in-sample dal 01/Gen/2008 al 31/Dic/2011. Pred.LL viene calcolata per il periodo out-of-sample dal 1/Gen/2012 al 31/Dic/2013. L'ultima colonna indica il numero di parametri di ogni modello.

2.3 Copulae e modelli GARCH

I modelli con Copulae esistenti in letteratura fanno un ampio uso dei modelli GARCH (Bollerslev, 1986) nella trattazione della volatilità delle singole serie. Nel primo step delle procedure descritte nei Paragrafi 2.1.2 e 2.1.5 vengono modellate le varianze condizionali di ciascuna serie storica attraverso dei modelli GARCH. La distribuzione multivariata dei residui standardizzati derivanti da tali modelli, viene analizzata nel secondo step, attraverso l'utilizzo delle copulae. A tal proposito, in questo paragrafo, viene fornita una breve rassegna sui modelli GARCH univariati esistenti e sulla loro generalizzazione nel multivariato.

2.3.1 Modelli GARCH univariati

Con l'introduzione dei modelli Auto-Regressive Conditional Heteroskedasticity (ARCH) Engle (1982) si pone l'attenzione sulla distinzione tra momento secondo condizionato e momento secondo non condizionato assumendo che i valori passati dello stesso abbiano una certa rilevanza. La dipendenza temporale diventa così un elemento di cruciale importanza. I modelli ARCH sono stati sviluppati per cogliere fenomeni di *volatility clustering* che caratterizzano

le distribuzioni delle serie storiche di variabili finanziarie.

Si assume che il set informativo al tempo t , I_t contenga tutte le informazioni storiche, fino a t , delle serie dei rendimenti.

Un modello di tipo ARCH identifica il termine di innovazione in una regressione lineare sui rendimenti y_t di una qualsiasi variabile finanziaria:

$$y_t = x_t' b + \epsilon_t$$

dove ϵ_t rappresenta l'innovazione al tempo t , x_t è il vettore di variabili esogene e b il vettore dei coefficienti ².

Si assume che:

$$E(\epsilon_t | I_{t-1}) = 0 \tag{2.3.1.1}$$

Questo implica che, per la legge dei valori attesi iterati, $E(\epsilon_t) = 0$ e quindi che tale processo non sia condizionatamente correlato al set informativo, infatti:

$$Cov(\epsilon_t \epsilon_{t+k} | I_{t-h}) = E(\epsilon_t \epsilon_{t+k} | I_{t-h}) - E(\epsilon_t | I_{t-h}) E(\epsilon_{t+k} | I_{t-h}) = 0$$

Il termine di innovazione ϵ_t è dato dalla relazione:

$$\epsilon_t = u_t h_t^{1/2} \tag{2.3.1.2}$$

dove $u_t \sim i.i.d.(0, 1)$ è un processo avente media nulla e varianza unitaria e h_t rappresenta la varianza condizionale dell'innovazione. In tale modello la varianza non condizionale risulta essere costante nel tempo, mentre la varianza condizionale potrebbe presentare una qualche variabilità temporale:

$$h_t = E(\epsilon_t^2 | I_{t-1}) = Var(\epsilon_t | I_{t-1}) \tag{2.3.1.3}$$

Secondo tale relazione il termine ϵ_t^2 viene per definizione assunto come un corretto previsore della volatilità.

²Se $b \neq 0$ la legge dei mercati efficienti fallisce

Date queste assunzioni abbiamo che l'innovazione si distribuisce condizionatamente come una Normale con media nulla e varianza pari ad h_t :

$$\epsilon_t | I_{t-1} \sim N(0, h_t) \quad (2.3.1.4)$$

Engle (1982) identifica la varianza condizionale in una funzione lineare nei parametri dei valori passati di ϵ_t^2 :

$$h_t = \alpha_0 + \sum_{i=1}^k \alpha_i \epsilon_{t-i}^2, \quad (2.3.1.5)$$

dove α_0 rappresenta il termine costante mentre α_i sono i coefficienti associati al termine di innovazione quadrato nel passato. L'equazione (2.3.1.5) rappresenta un caso particolare di un generico ARCH(k).

Condizione necessaria al fine di assicurare la positività della varianza è che i termini α_i ed α_0 sia tutti maggiori o uguali a zero. Riassumendo: un modello di regressione lineare su una serie storica di rendimenti y_t con innovazioni che seguono un modello di tipo ARCH può essere rappresentato nel seguente modo:

$$\left\{ \begin{array}{l} y_t = x_t' b + \epsilon_t \\ \epsilon_t = u_t h_t^{1/2} \\ h_t = E(\epsilon_t^2 | I_{t-1}) \\ y_t | I_{t-1} \sim N(x_t' b, h_t) \end{array} \right. \quad (2.3.1.6)$$

dove μ_t rappresenta la parte di media condizionale.

I modelli di tipo GARCH (*Generalized AutoRegressive Conditional Heteroskedasticity*), introdotti da Bollerslev (1986), si sono sviluppati nel campo di analisi delle serie storiche essendo processi non lineari in grado di spiegare il fenomeno della volatilità variabile.

Essi esprimono la varianza condizionale come funzione sia dei quadrati dei valori passati del termine dell'innovazione (ϵ_t), sia delle varianza condizionale

stessa:

$$h_t = \alpha_0 + A(L)\epsilon_t^2 + B(L)h_t \quad (2.3.1.7)$$

$$\eta_t = \epsilon_t^2 - h_t \quad (2.3.1.8)$$

dove $A(L)$ e $B(L)$ sono i polinomi nell'operatore ritardo rispettivamente di ϵ_t^2 e h_t .

Possiamo quindi facilmente ricavare che il modello GARCH è sostanzialmente un modello ARMA, infatti:

$$\epsilon_t^2 = \alpha_0 + A(L)\epsilon_t^2 + B(L)[\epsilon_t^2 - \eta_t] + \eta_t$$

ossia:

$$\{1 - [A(L) + B(L)]\}\epsilon_t^2 = \alpha_0 + [1 - B(L)]\eta_t \quad (2.3.1.9)$$

L'equazione (2.3.1.7) rappresenta un generico modello GARCH(p, q) dove q è l'ordine del polinomio $A(L)$ e p l'ordine del polinomio $B(L)$. La varianza condizionale, in questo caso, dipende da $(1 + p + q)$ parametri dove p si riferisce al numero dei ritardi della parte autoregressiva, mentre q fa riferimento all'ordine del ritardo del termine di errore (ϵ_t^2).

Quando le radici del polinomio $(1 - A(L) - B(L))$ cadono fuori dal cerchio unitario possiamo affermare che il modello è stazionario in covarianza. Viene da sé che quando, $p = 0$, il tutto si riduce ad un processo di tipo ARCH(q).

Stimare un processo di tipo GARCH(p, q) comporta necessariamente il fatto di dover stimare un numero $p + q + 1$ di parametri.

Il metodo di stima adottato è la Massima Verosimiglianza. Poiché la t -esima osservazione della variabile y_t non è indipendente dalle altre osservazioni, la funzione di verosimiglianza viene costruita come produttoria delle funzioni di densità condizionali (fattorizzazione sequenziale) piuttosto che delle funzioni di densità marginali. La procedura di massima verosimiglianza im-

pone che vengano fatte delle assunzioni sulla distribuzione delle innovazioni standardizzate

$$\frac{u_t = \epsilon_t}{\sqrt{h_t^{1/2}}}.$$

Distribuzioni come la t di Student e la Generalized Error Distribution sono state utilizzate come alternative alla più classica assunzione di normalità per cogliere la presenza di leptocurtosi. Rinunciare all'ipotesi di normalità dell'innovazione comporta l'utilizzo del metodo della quasi-massima verosimiglianza (QML). Il costo di questo passaggio è una stima meno efficiente di quella di massima verosimiglianza.

Esistono in letteratura svariate tipologie di modelli GARCH univariati (Bollerslev, Engle, and F. Nelson, 1994) come ad esempio:

- GARCH-M (GARCH in mean di Engle, Lilien, and Robins (1987))

$$\begin{cases} y_t = X_t b + \lambda g(h_t) + \epsilon_t \\ \epsilon_t | I_{t-1} \sim N(0, h_t) \end{cases} \quad (2.3.1.10)$$

dove $g(h_t)$ è una funzione continua e differenziabile che entra all'interno della funzione di regressione di y_t ;

- EGARCH (Exponential GARCH di Nelson (1990))

$$\log(h_t) = \omega + \sum_i^p \beta_i \log h_{t-j} + \sum_j^q \alpha_j g(u_t) \quad (2.3.1.11)$$

dove $g(u_t)$ è un processo *i.i.d.* avente media nulla.

La positività della varianza condizionale è assicurata dal logaritmo che rimuove tutti i vincoli di non negatività sui parametri.

L'EGARCH, attraverso la funzione $g(u_t)$, riesce a catturare l'asimmetria generata da uno shock sulla volatilità;

- GJR-GARCH (Glosten, Jagannathan, and Runkle, 1993)

$$h_t = \omega + \sum_{i=1}^q \{\alpha_i + \delta_i I(\epsilon_{t-i} > 0)\} \epsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} \quad (2.3.1.12)$$

dove I rappresenta la funzione indice. L'equazione (2.3.1.12) permette alla varianza condizionale di rispondere diversamente alle innovazioni passate, siano esse positive o negative.

- la sua versione non lineare STGARCH (Smooth Transition GARCH Gonzalo-Rivera (1998)) dove, attraverso una funzione di transizione G , viene reso possibile il passaggio da un regime all'altro.

$$h_t = \omega_1 + \sum_{i=1}^q \alpha_{1i} \epsilon_{t-i}^2 + (\omega_2 + \sum_{i=1}^q \alpha_{2i} \epsilon_{t-i}^2) G(\lambda, \gamma; \epsilon_{t-i}) + \sum_{i=1}^p \beta_i h_{t-i} \quad (2.3.1.13)$$

con

$$G(\lambda, \gamma; \epsilon_{t-i}) = (1 + e^{\lambda \prod_{j=1}^J (\epsilon_{t-i} - \gamma_j)})^{-1}, \quad \lambda > 0;$$

quando $J = 1$ tale funzione diviene una logistica.

- NGARCH (Non-linear GARCH introdotti da Higgins and Bera (1992))

$$h_t = \omega + \alpha (\epsilon_{t-1} - c \cdot h_{t-1}^{1/2})^2 + \beta h_{t-1} \quad (2.3.1.14)$$

dove $\omega > 0$, $\beta \geq 0$, $\alpha \geq 0$ e $\alpha(1+c^2) + \beta < 1$ assicurano la non negatività e la stazionarietà del processo;

- IGARCH (Integrated GARCH) (Baillie, Bollerslev, and Mikkelsen, 1996). Tale modello si ottiene quando il polinomio autoregressivo $A(L) + B(L)$ dell'equazione (2.3.1.7) assume valore unitario. In questo caso gli shock

sulla volatilità possiedono una forte persistenza;

- FIGARCH (Fractionally Integrated GARCH di Baillie, Bollerslev, and Mikkelsen (1996)) introdotto al fine di studiare quei processi che, pur non presentando una radice unitaria, non sono stazionari. In tale modello il polinomio nell'operatore ritardo viene definito come:

$$1 - A(L) - B(L) = f(L)(1 - L)^d$$

dove se $d = 1$ il modello corrisponde ad un IGARCH mentre se $d = 0$ siamo in presenza di un modello GARCH classico. ;

- TGARCH (Threshold GARCH di Zakoian (1994)) dove viene modellata la deviazione standar condizionale invece della varianza condizionale:

$$h_t^{1/2} = \omega + \sum_{i=1}^q (\beta_i h_{t-1}^{1/2} + \sum_{i=1}^q (\alpha_i^+ \epsilon_{t-i}^+ - \alpha_i^- \epsilon_{t-i}^-)) \quad (2.3.1.15)$$

con $\epsilon_t - i^+ = \max(\epsilon_t - i, 0)$ e $\epsilon_t - i^- = \min(\epsilon_{t-i}, 0)$.

2.3.2 GARCH Multivariati

In contesti di allocazione di portafoglio e risk hedging risulta più utile analizzare l'intera matrice di covarianza e per tale ragione sono stati sviluppati i modelli GARCH multivariati.

Nell'ambito multivariato abbiamo che y_t è un vettore composto da k serie ed ϵ_t è anch'esso un vettore con k termini di innovazione.

L'equazione 2.3.1.2 in questo caso diviene:

$$\epsilon_t = H_T^{1/2} u_t \quad (2.3.2.1)$$

dove H_t rappresenta la matrice delle covarianze condizionali e possiede una dimensione di $k \times k$ e u_t è un vettore di k elementi che soddisfa la condizione $E(u_t' u_t) = I_n$.

La generalizzazione di tali modelli può essere riassunta:

$$\begin{cases} y_t &= X_t B + \epsilon_t \\ \epsilon_t &= u_t H_t^{1/2} \\ u_t &\sim \text{i.i.d.}(0, I) \\ H_t &\sim \text{MGARCH} \end{cases} \quad (2.3.2.2)$$

Da questa rappresentazione si sono sviluppati diversi modelli GARCH multivariati dove l'elemento di differenziazione è dato dalla diversa specificazione della matrice H_t .

1. **VECH** (Bollerslev, Engle, and Wooldridge, 1988):

$$vech(H_t) = \gamma + \sum_{i=1}^q A_i vech(\epsilon_{t-i} \epsilon_{t-i}') + \sum_{j=1}^p B_j vech(H_{t-j})$$

dove l'operatore *vech* considera solo la porzione triangolare bassa di una matrice simmetrica ponendola all'interno di un vettore colonna. La specificazione di tale modello presenta due problemi: il primo è l'aumento del numero dei parametri da stimare al crescere del numero delle variabili considerate. Il secondo riguarda l'impossibilità di assicurare che la matrice H_t sia almeno semi-definita positiva anche imponendo dei vincoli sulle matrici A_i, B_j .

Bollerslev, Engle, and Wooldridge (1988) nel tentativo di ridurre il numero di parametri dell'intero modello, introducono una versione ristretta del VECH, il modello VECH diagonale. In questa semplificazione si assume che le matrici A_i e B_j siano diagonali. In tal modo il numero di parametri da stimare si riduce notevolmente ma la sua struttura risulta essere troppo restrittiva in quanto non sono ammesse interazioni tra le diverse covarianze e varianze condizionali;

2. **BEKK** (Engle and Kroner, 1995):

$$H_t = CC' + \sum_{i=1}^p \sum_{k=1}^K A'_{ki} \epsilon_{t-i} \epsilon'_{t-i} A_{ki} + \sum_{j=1}^q \sum_{k=1}^K B'_{kj} H_{t-j} B_{kj}$$

dove A_{ki} e B_{kj} sono matrici simmetriche non negative, C è una matrice triangolare alta e K sono i gradi di generalità. A differenza del modello VECH, in questo caso la matrice H_t è sempre definita positiva in quanto gli elementi che la compongono derivano dalla somma di forme quadratiche.

3. **GARCH fattoriali** (Engle, Ng., and Rothschild, 1990):

$$H_t = \Omega + \sum_{k=1}^K w_k w'_k f_{k,t}$$

dove Ω è una matrice semidefinita positiva, $f_{k,t}$ rappresentano i fattori e w_t il vettore dei pesi delle n attività rischiose. Questi modelli hanno avuto scarsa applicazione a livello empirico in quanto, essendo i fattori correlati tra loro, risulta molto difficoltoso effettuare una distinzione sugli effetti generati da ogni singolo fattore, non solo, risulta anche molto complesso individuare i fattori stessi.

4. **GARCH ortogonali** (Alexander and Chibumba, 1996), (Van Der Weide, 2002): dove le osservazioni originali sono funzione di una combinazione di fattori z_t non osservabili. Tali modelli si caratterizzano dalla seguente espressione:

$$y_t = W z_t$$

dove W è una matrice ortogonale ed invertibile ed i fattori z_t sono condizionatamente eteroschedastici.

5. **CCC** (Constant Conditional Correlation Bollerslev (1990))

Questo modello, a differenza di quelli appena visti, si concentra sulla specificazione delle varianze condizionali e le correlazioni. Constant

Conditional Correlation Bollerslev (1990)

$$H_t = D_t^{1/2} R D_t^{1/2}$$

dove D è una matrice diagonale composta dalle varianze condizionali ed R è la matrice delle correlazioni condizionali costanti nel tempo.

Il vantaggio di questo modello è di possedere una matrice H_t sicuramente definita positiva ed un numero di parametri da stimare inferiore rispetto ai modelli precedenti.

6. **DCC** (Dynamic Conditional Correlation (DCC), Engle (2002))

L'assunzione di correlazioni invarianti nel tempo porta però ad una scarsa adattabilità a livello empirico e per tale ragione sono stati sviluppati i modelli Dynamic Conditional Correlation (DCC), Engle (2002).

$$H_t = D_t^{1/2} R_t D_t^{1/2} \quad (2.3.2.3)$$

dove la matrice R_t diviene variante nel tempo. La stima di tale modello segue una procedura a due stadi. Il primo passo consiste nella stima delle varianze condizionali che seguono un GARCH(p,q) per poi passare alla stima della matrice R_t . I vantaggi che hanno portato questo modello ad essere ampiamente utilizzato nella letteratura finanziaria sono la riduzione del numero di parametri da stimare e la presa in considerazione di correlazioni variabili nel tempo. Lo svantaggio è che viene considerato un processo unico per l'intero insieme delle correlazioni condizionali.

I tradizionali modelli DCC (con innovazioni distribuite secondo una Normale ed una t) verranno utilizzati nel confronto delle performance con il modello che viene presentato nel secondo capitolo.

Capitolo 3

Modello di stima della struttura di dipendenza con copula multivariata dinamica

Lo scopo di questo capitolo è di introdurre un nuovo modello econometrico per la stima della struttura dipendenza esistente tra le serie dei rendimenti di diversi titoli azionari. Viene analizzato un nuovo modello con copula multivariata *Regular-vine* con parametri dinamici guidati da un modello *Generalized Autoregressive Score* (GAS). Le performance in-sample e le previsioni, ottenute con questo modello, vengono successivamente confrontate con modelli con copulae multivariata statica (R-vine, D-vine) e due modelli DCC.

3.1 Modello R-vine dinamico

Il modello presentato in questa sezione viene utilizzato per stimare la struttura di dipendenza esistente tra le serie dei rendimenti di attività finanziarie all'interno di un portafoglio di n titoli rischiosi. A tal fine, viene presa in considerazione la copula multivariata R-vine evitando di assumere vincoli restrittivi alla struttura di dipendenza come, invece, avviene nelle rappresentazioni D-vine e

C-vine (vedi Paragrafo 2.1.4), introducendo una dinamicità temporale nei parametri. Questo lavoro si inserisce all'interno del terzo filone di letteratura, descritto più nel dettaglio nel capitolo precedente, inerente lo studio di funzioni multivariate non normali attraverso l'utilizzo di modelli GARCH-Copula.

Dato un numero di serie storiche dei rendimenti > 2 , la procedura di stima prevede diversi passaggi riportati di seguito all'interno dell'Algoritmo 1:

Algorithm 1: Algoritmo di stima della copula multivariata R-vine dinamica

Input : n Serie storiche dei rendimenti;

Output: Matrice M , insieme dei parametri stimati;

- 1 Stima di un modello GARCH(1,1) per ogni singola serie storica e salvataggio dei residui standardizzati;
- 2 Trasformazione dei residui attraverso la funzione di distribuzione empirica (*empirical distribution function*);
- 3 Calcolo dell'indice di dipendenza $\tau_{j,k}$ di Kendall per ciascuna coppia di variabili $\{j, k\}$, $1 \leq j < k \leq n$;
- 4 Selezione della sequenza di variabili (*spanning tree*) che soddisfa le proprietà riportate nel Paragrafo 1.5.4 e che massimizza la somma dei valori assoluti dell'indice di Kendall:

$$\max_{\text{edges } e = (j,k) \text{ in spanning tree}} \sum |\tau_{j,k}|,$$

- 5 Per ogni *edge* $\{j, k\}$ all'interno della sequenza individuata, vengono stimate le copulae bivariate statiche e dinamiche e viene selezionata la copula che minimizza il Criterio di Informazione Bayesiano (BIC);
 - 6 Creazione delle pseudo-osservazioni (2.1.4.3)
 - 7 **for** $i = 2, \dots, n - 1$ (*iterazione sequenziale tra i diversi tree*) **do**
 - 8 Calcolo dell'indice di dipendenza $\tau_{i,j|D}$ per tutte le coppie $\{j, k|D\}$ appartenenti a T_i , cioè quelle coppie che soddisfano la condizione di prossimità (2.1);
 - 9 Ripetizione dei punti 4 e 5.
 - 10 **end**
-

La dinamicità temporale viene inserita, attraverso un modello Generalized Autoregressive Score, all'interno della struttura dei parametri.

Sia $\mathbf{Y}_t = [y_{1t}, y_{2t}]$ un insieme contenente le serie storiche dei rendimenti y_{1t}, y_{2t} con funzioni di distribuzione F_{1t} e F_{2t} , dato il set informativo $I_{t-1} =$

$(\mathbf{Y}_{t-1}, \mathbf{Y}_{t-2}, \dots)$, per il Teorema di Sklar, è possibile scrivere:

$$\mathbf{Y}_t | I_{t-1} \sim \mathbf{F}_t = C_t(F_{1t}(y_{1t}), F_{2t}(y_{2t}); \theta_{12,t})$$

dove \mathbf{F}_t è la funzione di distribuzione congiunta di y_{1t}, y_{2t} , C la funzione copula bivariata e θ_t il parametro time-varying della copula.

La formulazione appena riportata può essere riscritta come:

$$U_{it} = F_{it}(Y_{it}), \quad i = 1, 2$$

$$\mathbf{U}_t | I_{t-1} \sim C_t(u_{1t}, u_{2t})$$

dove $\mathbf{U}_t = [u_{1t}, u_{2t}]$

Si assume che il parametro time-varying sia caratterizzato da un modello di tipo GAS (Creal, Koopman, and Lucas, 2013):

$$\begin{cases} \theta_t &= \omega + \phi\theta_{t-1} + \alpha s_{t-1} \\ s_{t-1} &= S_{t-1} \nabla_{t-1} \\ \nabla_{t-1} &= \frac{\partial \ln c(\mathbf{u}_t; \theta_{t-1})}{\partial \theta_{t-1}} \end{cases} \quad (3.1.0.1)$$

con S_t matrice di scala rappresentata dall'inversa della matrice Hessiana o dalla sua radice quadrata. Anche se tale specificazione della dinamica time-varying del parametro può sembrare del tutto arbitraria, Creal, Koopman, and Lucas (2013) motivano tale metodo dimostrando come esso possa comprendere diversi modelli molto noti in letteratura ed ampiamente utilizzati come ad esempio i modelli GARCH per la varianza condizionata (Bollerslev, 1986) ed i modelli di trade durations (ACD)(Engle and Russell, 1998).

Combinando il modello con copula bivariata dinamica appena riportato con la metodologia R-vine (si veda il Paragrafo 1.5.4), otteniamo la seguente

funzione di densità della copula multivariata dinamica:

$$c(u_1, \dots, u_n; \theta_t) = \prod_{i=1}^{n-1} \prod_{e \in E_i} c_{j(e),k(e)|D(e)}(F(u_{j(e)}|\mathbf{u}_{D(e)}), F(u_{k(e)}|\mathbf{u}_{D(e)})), \quad (3.1.0.2)$$

con $\mathbf{u} = (u_1, \dots, u_n)' \in [0, 1]^n$ e $\theta_t = \{\theta_{t;j(e),k(e)|D(e):e \in E_i, i=1, \dots, n}\}$ dove $\theta_{j(e),k(e)|D(e)}$ indica il parametro della copula bivariata $c_{j(e),k(e)|D(e)}$. E_i rappresenta l'insieme di tutte le *edge* della rappresentazione R-vine, D è l'insieme degli elementi condizionanti e $\mathbf{u}_{D(e)}$ il vettore \mathbf{u} degli elementi contenuti all'interno di $D(e)$.

Dal Teorema di Sklar abbiamo che la funzione di densità di una qualsiasi distribuzione multivariata è composta dalla produttoria delle densità marginali moltiplicata per la densità della copula multivariata:

$$f(y_1, \dots, y_n) = \left[\prod_{i=1}^n f_i(y_i) \right] \times c_{1\dots n}(F_1(y_1), \dots, F_d(y_n)) \quad (3.1.0.3)$$

Data la matrice R-vine $M = (m_{ij}|i, j = 1, \dots, n)$ (Paragrafo 1.5.4), in conformità con Dißmann, Brechmann, Czado, and Kurowicka (2013), la funzione di densità (3.1.0.3) diviene:

$$f(y_1, \dots, y_n) = \left[\prod_{i=1}^n f_i(y_i) \right] \times \left[\prod_{j=n-1}^1 \prod_{k=n}^{j+1} c_{m_{j,j}, m_{k,j}|m_{k+1,j}, \dots, m_{n,j}} \right] \quad (3.1.0.4)$$

dove gli argomenti delle copulae bivariate sono stati omessi per semplicità di notazione.

La struttura sequenziale della copula R-vine viene costruita attraverso l'utilizzo l'algoritmo *maximum spanning tree* (punti 3 e 4 dell'Algoritmo 1), che massimizza la somma dei valori assoluti degli indici di Kendall delle diverse variabili. Tale algoritmo opera in maniera analoga all'algoritmo *minimum spanning tree* proposto da Prim (Cormen, Leiserson, Rivest, and Stein, 2001). I tree creati con tale metodo sono invarianti rispetto a trasformazioni monotone crescenti dei valori delle serie utilizzate.

3.2 Stima del modello R-vine dinamico

In questa sezione viene descritta la procedura di stima dell'intero modello proposto. Viene utilizzata una procedura semiparametrica come quella descritta Paragrafo 2.1.2.

Inizialmente viene stimato un modello GARCH(1,1) per ogni singola serie storica attraverso la massimizzazione della funzione di verosimiglianza. La funzione di distribuzione congiunta dei residui standardizzati estratti da tale modello viene trattata con il modello R-vine dinamico. A tal proposito, i residui standardizzati vengono trasformati utilizzando la funzione di distribuzione empirica (*empirical cumulative distribution function*) in modo da ottenere valori $\in [0, 1]$. L'applicazione di tale trasformazione non comporta alcuna alterazione alla struttura di dipendenza esistente. Seguendo questa procedura possiamo stimare separatamente i parametri delle distribuzioni marginali e della copula. La copula multivariata R-vine, come le altre vine-copula, viene costruita come prodotto di copulae bivariate (PCC). Ciò comporta un ulteriore vantaggio, in quanto è possibile stimare i parametri della copula in maniera sequenziale piuttosto che in unico passaggio, riducendo così la complessità dell'intera procedura ed i tempi di macchina.

A differenza del modello R-vine statico in questo caso abbiamo bisogno delle funzioni *score* (∇_{t-1}) di ogni singola copula bivariate nonché della matrice di scala S_{t-1} . In questo modello vengono analizzate le versioni dinamiche delle copulae bivariate descritte nel Paragrafo 1.5.1.

Di seguito vengono fornite le funzioni score delle log-densità di ciascuna copula:

1. Copula Gaussiana:

$$\nabla_N = \frac{1}{c_N} \times \left(- \frac{(\rho^3 - x_1 x_2 \rho^2 + \rho(x_1^2 + x_2^2 - 1) - x_1 x_2) \exp \left\{ \frac{\rho^2(x_1^2 + x_2^2) - 2\rho x_1 x_2}{2(\rho^2 - 1)} \right\}}{(1 - \rho^2)^{\frac{5}{2}}} \right) \quad (3.2.0.1)$$

dove $x_1 = \Phi^{-1}(u_1)$ e $x_2 = \Phi^{-1}(u_2)$ e c_N è la densità della copula Gaussiana.

2. t Copula:

$$\left\{ \begin{array}{l} \nabla_{t(\rho)} = -(v+1)\frac{\rho}{1-\rho^2} + (v+2)\frac{v\rho + x_1x_2}{M(v, \rho, x_1, x_2)} \\ \nabla_{t(v)} = -\psi\left(\frac{v+1}{2}\right) + \psi\left(\frac{v}{2}\right) + \frac{1}{2}\ln(1-\rho^2) - \frac{v-2}{2v} - \frac{1}{2}\ln(v) \\ \quad \frac{v+1}{2} \left[\frac{1+2x_1+x_1^2G(x_1)}{v+x_1} + \frac{1+2x_2G(x_2)}{v+x_2^2} \right] \\ \quad - \frac{v+2}{2} \frac{1-\rho^2 + 2x_1G(x_1) + 2x_2G(x_2) - 2\rho(x_1G(x_2) + x_2G(x_1))}{M(v, \rho, x_1, x_2)} \\ \quad + \frac{1}{2}(\ln((v+x_1^2)(v+x_2^2)) - \ln(M(v, \rho, x_1, x_2))) \end{array} \right. \quad (3.2.0.2)$$

dove $G(\cdot) = \frac{\partial(\cdot)}{\partial v}$, $\psi(\cdot)$ è la funzione digamma e $M(v, \rho, x_1, x_2) = v(1 - \rho^2) + x_1^2 + x_2^2 - 2\rho x_1x_2$.

3. Frank Copula:

$$\nabla_F = \left(\frac{1}{\theta} + e^{-\theta}(1 - e^{-\theta})^{-1} - (u_1 + u_2) - 2[(1 - e^{-\theta}) - t_1t_2]^{-1}(e^{-\theta} - \frac{\partial t_1}{\partial \theta}t_2 + t_1\frac{\partial t_2}{\partial \theta}) \right) \quad (3.2.0.3)$$

dove

$$t_i = (1 - e^{-\theta u_i}), \quad i = 1, 2$$

4. Clayton Copula:

$$\nabla_C = \left(\frac{\ln A(u_1, u_2, \theta)}{\theta^2} + \frac{(-2 - \frac{1}{\theta})(-u_1^{-\theta} \ln u_1 - u_2^{-\theta} \ln u_2)}{A(u_1, u_2, \theta)} \right) - \ln u_1u_2 \quad (3.2.0.4)$$

con $A(u_1, u_2, \theta) = (u_1^\theta + u_2^{-\theta} - 1)$.

5. Gumbel Copula:

$$\begin{aligned}
\nabla_G = & \left[-g^{\frac{1}{\theta}} \left(-\frac{\ln(g)}{\theta^2} + \frac{t_1 \ln(-\ln(u_1)) + t_2 \ln(-\ln(u_2))}{\theta(g)} \right) \right. \\
& \left(-2\frac{\ln(g)}{\theta^2} + \left(-2 + \frac{2}{\theta} \right) \frac{t_1 \ln(-\ln(u_1)) + t_2 \ln(-\ln(u_2))}{g} \right) + \ln(\ln(u_1) \ln(u_2)) \Big] \\
& + C_G(u_1, u_2) g^{-2 + \frac{2}{\theta}} \frac{(\ln(u_1) \ln(u_2))^{\theta-1}}{u_1 u_2} \\
& \left((g)^{-\frac{1}{\theta}} + (\theta - 1)(g)^{-\frac{1}{\theta}} \left(\frac{\ln(g)}{\theta^2} - \frac{t_1 \ln(-\ln(u_1)) + t_2 \ln(-\ln(u_2))}{\theta g} \right) \right) \times \frac{1}{c_G}
\end{aligned} \tag{3.2.0.5}$$

dove $t_i = (-\ln(u_i))^\theta$ con $i = 1, 2$ e $\theta \geq 1$ e $g = \sum_{i=1}^2 t_i$. C_G è la funzione di distribuzione della copula Gumbel e c_G la sua funzione di densità.

Come visto nel sistema (3.1.0.1) la funzione score della copula viene moltiplicata per una matrice di scala S predefinita, andando così a formare la funzione *scaled score*. In questo lavoro viene utilizzata l'inversa della radice quadrata della matrice Hessiana come matrice di scala e, per non appesantire ulteriormente la notazione non vengono riportate le funzioni analitiche delle derivate seconde.

Dopo aver ricavato le diverse funzioni *scaled-score* per ogni copula, la stima viene condotta attraverso la massimizzazione numerica della funzione log-verosimiglianza di ciascuna copula bivariata dove l'evoluzione del parametro viene dettata da un modello GAS (3.1.0.1). Come è possibile notare, l'introduzione di una dinamicità nei parametri attraverso l'utilizzo di un modello GAS all'interno del meccanismo PCC, comporta un aumento dei parametri da stimare. Per ogni copula bivariata dinamica considerata, infatti, il numero di parametri da stimare è pari a 3 o 4 nel caso di copula t-Student, mentre nel caso di copula bivariata statica è pari ad 1 o 2 (copula t-Student). Tale complicazione comporta inevitabilmente un aumento nei tempi-macchina di stima dell'intero modello (Tabella 3.3.3). Nel tentativo di ridurre la complessità della stima, viene analizzata la possibilità di troncamento proposta da Brechmann, Czado, and Aas (2012) (Paragrafo 1.8).

Nella selezione del livello di troncamento viene utilizzato il test di rapporto di verosimiglianza proposto da Vuong (1989).

Al fine di comparare due differenti modelli, f_1 e f_2 , (in questo caso la differenza tra i due modelli è data dal livello di troncamento), con parametri stimati $\hat{\theta}_1$ e $\hat{\theta}_2$, viene calcolata la somma standardizzata, s , dei valori $m_i = \log \left[\frac{f_1(x_i|\hat{\theta}_1)}{f_2(x_i|\hat{\theta}_2)} \right]$ della stima di verosimiglianza per ogni osservazione x_i con $i = 1, \dots, n$. Sotto le condizioni generali di regolarità, s si distribuisce asintoticamente come una Normale standardizzata.

Il modello 1 viene preferito al modello 2 se:

$$s = \frac{\frac{1}{n} \sum_{i=1}^n m_i}{\sqrt{\sum_{i=1}^n (m_i - \bar{m})^2}} \geq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \quad (3.2.0.6)$$

dove Φ^{-1} rappresenta la funzione inversa della distribuzione normale standardizzata. Se $s < -\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$ viene preferito il modello 2, mentre, se $|s| \leq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$ non si è in grado di esprimere delle preferenze tra i due modelli. L'ipotesi nulla del test è l'uguaglianza tra il modello senza troncamento ed il modello con un livello di troncamento K . All'interno dell'Algoritmo 2 viene riportata una sintesi di questa procedura.

Algorithm 2: Selezione del livello di troncamento K

Input : n Serie storiche dei rendimenti, livello di confidenza α ;
Output: Copula R-vine troncata;

- 1 **for** $i = 0, \dots, n - 2$ **do**
- 2 Costruzione della copula multivariata $tRV(i + 1)$
- 3 Calcolo della statistica test s (3.2.0.6) dove il modello f_1 è rappresentato dalla copula $tRV(i)$ e f_2 dalla copula $tRV(i + 1)$
- 4 **if** $s \geq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$ **then**
- 5 Troncare la copula R-vine a livello $K = i$ ed uscire dal loop
- 6 **end**
- 7 **end**

La funzione di densità della copula R-vine dinamica troncata è del tutto

simile all'equazione (2.2.3.1):

$$c_{tRV(K)}(y_1, \dots, y_n | \boldsymbol{\theta}_{tRV(K)}^t) = \prod_{i=1}^K \prod_{e \in E_i} c_{j(e), k(e) | D(e)}(F(y_{j(e)} | \mathbf{y}_{D(e)}), F(y_{k(e)} | \mathbf{y}_{D(e)})), \quad (3.2.0.7)$$

dove però il parametro $\boldsymbol{\theta}$ è variante nel tempo. La densità (3.2.0.7) può essere intesa anche come una *composite likelihood* (Lindsay, 1988).

Consideriamo un vettore n -dimensionale Y , con funzione di densità $f(y; \theta)$ ed indichiamo con $\{\mathcal{A}_1, \dots, \mathcal{A}_K\}$ l'insieme delle marginali o degli eventi condizionanti la cui funzione di verosimiglianza $L_k(y; \theta) \propto f(y \in \mathcal{A}_k; \theta)$. La funzione *composite likelihood* (Lindsay, 1988) è data da:

$$L_C(y; \theta) = \prod_{k=1}^K L_k(y; \theta)^{w_k} \quad (3.2.0.8)$$

dove w_k sono dei pesi non negativi che devo essere scelti. A differenza di quanto avviene nella metodologia *composite likelihood*, nel modello R-vine dinamico troncato non vi è la necessità di scegliere alcun peso da inserire nell'equazione (3.2.0.10) restando comunque una valida funzione di densità.

Lo stimatore $\hat{\theta}_{CL}$ della funzione 3.2.0.8, definito come *maximum composite likelihood estimator*, rappresenta il punto di massimo della funzione 3.2.0.8 o equivalentemente della funzione $cl(y; \theta) = \sum_{k=1}^K l_k(y; \theta)_{w_k}$, dove $l_k(y; \theta) = \log L_k(y; \theta)$. Lo stimatore $\hat{\theta}_{CL}$ viene ricavato risolvendo la funzione *composite score* $c_s(y; \theta) = \nabla_{\theta} cl(y; \theta)$ rappresentata dalla combinazione lineare degli score di ciascun termine $l_k(y; \theta)$.

In presenza di n osservazioni indipendenti ed identicamente distribuite Y_1, \dots, Y_n ricavate dal modello $f(y, \theta)$, lo stimatore $\hat{\theta}_{CL}$ risulta essere consistente ed asintoticamente normale:

$$\sqrt{n} (\hat{\theta}_{CL} - \theta) \xrightarrow{d} N(0, G^{-1}(\theta))$$

dove G è la matrice di informazione di Godambe (1960):

$$G(\theta) = H(\theta) J(\theta)^{-1} H(\theta)$$

Le matrici H e J rappresentano rispettivamente la *sensitivity matrix* e la *variability matrix*:

$$H(\theta) = \mathbb{E}_\theta\{-\nabla_\theta c_s(Y; \theta)\} = \int \{-\nabla_\theta c_s(y; \theta)\} f(y; \theta) dy$$

$$J(\theta) = \text{var}_\theta\{c_s(Y; \theta)\}$$

Ragionando in termini di *composite likelihood* la funzione di densità (3.1.0.4) può essere scomposta anche nel seguente modo:

$$f(y_1, \dots, y_n) = f_{m_{n,n}}(y_{m_{n,n}}) \times \prod_{i=2}^n f(y_{m_{n-i+1, n-i+1}} | y_{m_{n-i, n-i+1}}, \dots, y_{m_{n, n-i+1}}), \quad (3.2.0.9)$$

dove la sequenza delle variabili condizionali è data dagli elementi in colonna di una matrice R-vine M già definita (Paragrafo 1.5.4). Assumendo un livello di troncamento pari a K , la funzione di densità diviene:

$$f(y_1, \dots, y_n) = f_{m_{n,n}}(y_{m_{n,n}}) \times \prod_{i=2}^n f(y_{m_{n-i+1, n-i+1}} | y_{m_{n-K+1, n-i+1}}, \dots, y_{m_{n, n-i+1}}), \quad (3.2.0.10)$$

dove le copulae bivariate appartenenti a T_i con $K + 1 \neq i \neq n - 1$ vengono impostate come *independence copula*. Con un livello $K = 1$ si ottiene un modello Markov-chain di ordine 1, mentre ponendo che tutte le distribuzioni marginali e tutte le copulae bivariate seguono una distribuzione Normale allora l'equazione 3.2.0.10 coincide con il modello introdotto da Vecchia (1988).

Vista l'analogia di tali modelli con il modello R-vine dinamico possiamo supporre che la stima di $\theta_{tRV(K)}^t$ nell'equazione (3.2.0.7) sia anch'essa consistente (Lindsay (1988), Creal, Koopman, and Lucas (2013), Eike, Brechmann,

and Czado (2013), Almeida, Czado, and Manner (2016))

3.3 Valutazione delle performance

In questo paragrafo vengono valutate le performance in-sample ed out-of-sample del modello proposto su un dataset di 17 serie storiche dei rendimenti (la lista delle serie utilizzate e le statistiche descrittive vengono riportate in Appendice).

Sono stati utilizzati dati giornalieri di titoli azionari quotati all'interno del FTSE-MIB per il periodo dal 01/04/2010 al 22/03/2017 per un totale di 1775 osservazioni. Le prime 1400 osservazioni (circa 2/3 dell'intero dataset) sono state utilizzate nell'analisi in-sample mentre le restanti 375 per le valutazioni delle performance previsive.

Il periodo considerato ha mostrato una forte instabilità ed altalenanza nelle performance dell'indice. Negli anni 2010 e 2011 si sono registrati dei decrementi rispettivamente del $-13,23\%$ e del $-25,2\%$, quest'ultima risultata essere una delle peggiori performance tra le borse internazionali. La differenza percentuale tra valore massimo e minimo dell'anno 2011 è stata di circa il 72% , segno, appunto, di una forte instabilità dell'intero mercato. Gli anni 2012 e 2013 hanno segnato una ripresa registrando incrementi del $+7,8\%$ e del $+16,56\%$. Nonostante la buona performance del 2012, alcuni titoli, come ad esempio Mediaset e Telecom Italia, considerati all'interno dell'analisi, hanno comunque riportato dei forti ribassi rispettivamente del $27,2\%$ e del $17,8\%$. L'anno 2014 ha registrato una situazione pressoché invariata con un $+0,23\%$ mentre nei successivi due anni vi è stata un'inversione di tendenza con delle performance rispettivamente del $-5,72\%$ e del $-10,2\%$. Anche nel primo trimestre del 2017 si è registrato un incremento nell'andamento dell'indice. Periodi di forte instabilità ed incertezza nei mercati generano un incremento nella rischiosità degli investimenti finanziari, pertanto, è di fondamentale importanza per il

portfolio manager avere a disposizione degli strumenti che possano al meglio gestire tali situazioni rischiose. A tal proposito, in questo lavoro si è voluto indagare sull'utilizzo di un modello con copula multivariata R-vine dinamica all'interno di problematiche di allocazione di portafoglio come alternativa ai modelli classici. Sono stati considerati solo 17 titoli azionari facenti parti del paniere FTSE-MIB. La scelta è dovuta alla complessità del modello ed ai tempi macchina di stima, che aumenterebbero all'aumentare del numero di variabili considerate.

Le serie dei rendimenti $r_{i,t}$ sono state ricavate in regime di capitalizzazione continua:

$$r_{i,t} = 100 \times [\ln(P_{i,t}) - \ln(P_{i,t-1})]$$

dove $P_{i,t}$ rappresenta la quotazione dell' i -esimo titolo.

Sono state inizialmente modellate le varianze condizionali attraverso un modello GARCH(1,1) con errori estratti da una distribuzione t , in modo da tener conto degli eventi estremi. La soluzione con errori estratti da una distribuzione *skewed-t* non viene trattata per evitare di complicare ulteriormente la procedura di stima e simulazione, ma sarà oggetto di lavori futuri. Per la struttura di dipendenza sono stati messi a confronto 6 diversi modelli. Il primo è un modello full R-vine dinamico dove non viene considerata la possibilità di troncamento, il secondo è un modello R-vine dinamico troncato dove il livello di troncamento K è stato ricavato secondo la procedura utilizzata da Brechmann, Czado, and Aas (2012) descritta nel dettaglio all'interno del Paragrafo 1.8 e nell'Algoritmo 2. Questi modelli sono poi stati confrontati con un modello R-vine statico, un modello D-vine statico, un DCC con innovazioni Gaussiane ed un DCC con innovazioni estratte da una distribuzione t .

All'interno della procedura di stima sono state utilizzate 14 diverse copulae bivariate: la copula Gaussiana (N), la copula t-Student (t), la copula Clayton (C), la copula Gumbel (G), la copula Frank (F) considerando le versioni statiche e dinamiche. Sono anche state valutate le versioni, statiche e dinami-

che delle copulae Clayton (RC) e Gumbel (RG) ruotate di 180° . All'interno della copula t-Student dinamica viene considerato variabile nel tempo solo il parametro di correlazione ρ rimanendo invece costante il parametro relativo ai gradi di libertà.

La scelta di quale copula bivariata utilizzare per ogni singola coppia viene effettuata attraverso la comparazione del criterio di informazione Bayesiano (BIC) mentre per la valutazione del livello di troncamento invece, viene utilizzato il test di rapporto di verosimiglianza proposto da Vuong (1989) (3.2.0.6).

3.3.1 Stime e performance in-sample

In questa sezione vengono mostrati i risultati ottenuti all'interno della porzione in-sample del campione analizzato. Nella tabella 3.3.1 sono riportati i valori dettagliati delle stime del primo *tree* del modello R-vine dinamico. Si può notare che il meccanismo di selezione individua 9 copulae dinamiche e 7 statiche. I valori elevati del parametro α validano la presenza di una dinamica temporale nell'evoluzione dei rapporti di dipendenza. La scelta tra le copulae dinamiche ricade spesso sulla copula t indicando come la distribuzione della struttura di dipendenza sia leptocurtica. I valori dettagliati delle stime dei successivi *tree* non vengono riportati per brevità di notazione.

Data l'elevata complessità del modello con 17 variabili viene considerata la possibilità di troncamento utilizzando il test di rapporto di verosimiglianza proposto da Vuong (1989) (3.2.0.6). Nella Tabella 3.3.2 vengono riportanti i valori di log-verosimiglianza, del test di Vuong e del criterio informativo bayesiano (BIC). Si nota subito che, anche se il modello R-vine dinamico non troncato ottiene i più alti valori di verosimiglianza, non avviene lo stesso per il criterio informativo BIC. Tale risultato è dovuto probabilmente ad una sovra-parametrizzazione di tale modello. Analizzando le statistiche test riportate nella terza colonna si osserva che è possibile troncamento il modello

Tabella 3.3.1: Stima dei parametri del primo tree del modello R-vine dinamico nella sezione in-sample

tree	edge	Copula			GAS		
		family	par	par2	omega	phi	alpha
1	9,10	D-t	-	6.44	0.81	0.19	0.921
	8,9	D-t	-	4.31	0.77	0.06	0.903
	3,9	t	0.75	5.59	-	-	-
	6,8	t	0.69	5.86	-	-	-
	2,9	D-t	-	10.76	0.68	0.11	0.911
	6,16	t	0.65	7.67	-	-	-
	8,12	D-RG	-	-	2.28	0.15	0.955
	9,11	D-RG	-	-	1.67	0.18	0.923
	4,12	t	0.61	8.97	-	-	-
	1,8	D-RG	-	-	1.79	0.03	0.981
	7,9	D-t	-	7.59	0.57	0.10	0.927
	8,17	t	0.58	7.03	-	-	-
	6,14	D-t	-	5.90	0.62	0.05	0.899
	12,15	t	0.53	8.08	-	-	-
	5,8	D-N	-	-	0.37	0.09	0.894
	1,13	G	1.31	-	-	-	-

La colonna *family* indica la famiglia che è stata selezionata per la coppia di variabili indicata nella colonna *edge*. I termini D-t, D-RG e D-N rappresentano rispettivamente la copula bivariata t di Student dinamica, la copula di Gumbel ruotata dinamica e la copula Gaussiana dinamica. Il parametro di correlazione ρ della copula t-Student dinamica corrisponde al parametro omega della stima GAS.

dinamico a livello $K = 6$ lasciando quindi come implicitamente indipendenti tutte le copulae dei successivi tree. Il medesimo risultato viene raggiunto anche dal confronto tra i diversi valori del criterio informativo bayesiano dei diversi modelli. Ciò comporta una riduzione sia del numero di parametri da stimare, sia nei tempi macchina di stima, come evidenziato all'interno delle ultime due colonne della Tabella 3.3.2, rendendo il modello dinamico di più facile applicazione. Nei modelli R-Vine-st e D-Vine vengono utilizzate solo le versioni statiche delle copulae Gaussiana, t, Frank, Clayton, Rotated Clayton, Gumbel, Rotated Gumbel. All'interno del dataset considerato, questi modelli ottengono le peggiori performance in assoluto evidenziando così la loro scarsa

Tabella 3.3.2: Confronto performance in-sample

Modello	Log-lik.	Vuong test	BIC	n. par	tempo di stima
R-vine-D	12516.84	-	-25767.36	231	1451 sec.
Tr. R-vine-D (6)	12295.49	0.82*	-25792.19	148	843 sec.
R-vine-st	9321.51	3.24	-17472.91	160	215 sec.
D-vine model	7782.01	4.18	-14396.91	160	150 sec.
DCC -G	9948.06	4.09	-19354.94	74	96 sec.
DCC - t	10054.26	4.01	-19561.56	91	101 sec.

Il test riportato nella colonna 3 mette a confronto, ad un livello $\alpha = 0.05$, i modelli considerati con il modello R-Vine-D. Il simbolo “*” accanto alla statistica test indica che il modello considerato è indistinguibile o anche superiore al modello R-Vine-D.

Tabella 3.3.3: Copulae selezionate all’interno di un modello R-vine dinamico troncato a livello 6

Tree	time-varying							time-constant						
	N	t	C	RC	F	G	RG	N	t	C	RC	F	G	RG
1	1	5	-	-	-	-	3	-	6	-	-	-	1	-
2	1	4	-	-	-	-	1	-	2	-	-	5	2	-
3	-	2	-	-	-	-	1	2	-	1	1	7	-	-
4	-	1	-	-	-	-	-	-	-	2	-	5	1	4
5	-	-	-	-	-	-	-	-	2	-	3	6	1	-
6	-	-	-	-	-	-	-	2	-	3	-	4	2	-

utilità nel contesto analizzato.

Dalla tabella 3.3.2 emerge una possibile non linearità nella relazione tra la performance e la complessità del modello: il passaggio da un meno complesso DCC ad un più complesso modello con copulae statiche peggiora la performance del portafoglio, ma un ulteriore aumento della complessità tramite l'uso di copulae dinamiche inverte questo risultato. Tale non linearità è evidenziata anche nella procedura di troncamento dove una riduzione della complessità del modello conduce alle medesime, o anche migliori, performance di portafoglio del modello completo.

Nella Tabella 3.3.3 vengono riportate le tipologie ed il numero di copulae bivariate selezionate all'interno del modello R-vine dinamico troncato. I risultati mostrano che vi è una diminuzione nel numero di copulae dinamiche che vengono selezionate indicando che la dinamicità temporale viene interamente catturata, come da attese, all'interno dei primi 4 *tree*. Nei modelli con copula multivariata, costruiti attraverso il meccanismo del *Pair Copula Construction* vengono modellati, all'interno dei primi *tree*, quei rapporti di dipendenza più forti che influenzano maggiormente le performance di portafoglio. Le scelte effettuate in questi primi *tree* andranno poi ad influenzare l'intera struttura sequenziale, pertanto, riuscire ad individuare e catturare una dinamicità temporale all'interno della struttura di dipendenza, risulta essere di fondamentale importanza.

I risultati ottenuti da questa prima analisi suggeriscono che è possibile modellare un dataset anche di più grandi dimensioni in quanto la struttura di dipendenza maggiormente rilevante viene catturata all'interno dei primi *tree*. Risulta inoltre che, in un contesto di forte instabilità del mercato come quello analizzato, il modello R-vine dinamico introdotto in questo lavoro riesce meglio ad adattarsi ai dati ottenendo le migliori performance in assoluto. Ciò dimostra come uno studio più approfondito e dettagliato dei comportamenti congiunti delle serie dei rendimenti, effettuato attraverso l'analisi della struttura di dipendenza, possa portare risultati migliori rispetto ai modelli classici.

Questo risultato ha un grande valore all'interno delle strategie di allocazione di portafoglio. Un'analisi in merito viene proposta nel Paragrafo 4.2.

3.3.2 Previsioni

Al fine della validazione del modello vengono analizzate e confrontate le capacità previsionive. Una volta stimato il modello R-vine dinamico, i parametri ottenuti vengono utilizzati per effettuare delle previsioni ad un passo. Come in Almeida, Czado, and Manner (2016) le capacità previsionive vengono indagate considerando diversi criteri. Il primo riguarda la log-likelihood prevista, mentre il secondo considera due tipologie di test retrospettivi (*backtesting*).

La log-likelihood prevista, riportata nella Tabella 3.3.4, viene calcolata all'interno della porzione out-of-sample del campione. In altre parole, il modello stimato nella parte in-sample del campione viene applicato ai dati out-of-sample. I risultati mostrano che, anche in questo caso, i modelli con copula multivariata dinamica raggiungono performance migliori rispetto agli altri modelli considerati. Nella Tabella 3.3.5 vengono invece riportati i p-value dei due diversi test retrospettivi ovvero il test *Proportion Of Failure* introdotto da Kupiec (1995) conosciuto anche come test *unconditional coverage* delle previsioni del VaR ed il test *conditional coverage* proposto da Christoffersen (1998).

Test di Kupiec (1995)

Il test proposto da Kupiec (1995), anche definito come *Proportion of Failure test*, concentra l'attenzione sul numero di perdite giornaliere eccedenti il VaR stimato. Dato un numero T di perdite osservate superiori al VaR, calcolato ad un livello di confidenza α , il test di Kupiec risponde alla domanda se il rapporto T/N , con N numero di osservazioni totali, sia statisticamente differente da α . L'ipotesi nulla è che il numero di eccezioni rilevate empiricamente sia coerente con in numero di eccezioni attese. Tale verifica non è condizionata a ulteriori ipotesi, perciò il test è *unconditional*.

La probabilità di osservare un numero di perdite T su un numero totale di osservazioni N , segue la distribuzione Binomiale $(1 - \alpha)^{N-T} \alpha^T$. Il test per

l'ipotesi nulla $T/N = \alpha$ si basa sul rapporto di verosimiglianza ed è dato da:

$$LR_{UC} = 2 \ln[(1 - T/N)^{N-T} (T/N)^T] - 2 \ln[(1 - \alpha)^{(N-T)} \alpha^T] \quad (3.3.2.1)$$

che, sotto H_0 si distribuisce come una $\chi^2(1)$. Tale test difetta di scarsa potenza statistica in quanto, per generare dei dati veramente affidabili, necessita di un numero di osservazioni molto elevato.

Test di Christoffersen (1998)

Christoffersen propone un test più completo del precedente in cui viene posta a verifica congiunta l'ipotesi di *unconditional coverage*' (Kupiec 1995) e l'ipotesi di indipendenza delle perdite in eccesso. La differenza principale è che questo è un test di tipo *conditional*: ad esempio se la volatilità risulta essere bassa in determinati periodi ed alta in altri, il VaR previsto deve tenere conto di questi *cluster*.

La statistica test è rappresentata da:

$$LR_{CC} = 2 \ln[(1 - \pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1 - \pi_{11})^{n_{10}} \pi_{11}^{n_{11}}] - 2 \ln[(1 - \alpha)^{N-T} \alpha^T] \quad (3.3.2.2)$$

dove π è dato da:

$$\pi_{ij} = \frac{n_{ij}}{\sum_j n_{ij}}$$

dove n_{ij} è il numero di osservazioni che presentano il valore i seguito dal valore j con $i, j = 0, 1$.

La statistica test si distribuisce, sotto H_0 , come una $\chi^2(2)$. Se la sequenza T delle perdite in eccesso rispetto VaR è indipendente, allora $\pi_{01} = \pi_{11} = \alpha$. Attraverso questa statistica è possibile testare unicamente l'indipendenza delle perdite in eccesso, non la correttezza del modello.

3.3.2.1 Risultati

Dai valori di log-likelihood prevista si nota come i due modelli con copula R-vine dinamica ottengono i migliori risultati. Inoltre, come nell'analisi in-sample, i risultati mostrano che non è possibile distinguere il modello troncato dal modello non troncato. Come già evidenziato in precedenza, vengono utilizzate 375 osservazioni per valutare la bontà previsionale del modello. Per livelli di confidenza pari al 10%, 5% e 1% ci si aspetta che la perdita massima venga superata rispettivamente un numero di volte pari a 37.5, 18.75 e 3.75. Nella Tabella 3.3.5 risulta evidente che nel caso dei modelli con copula dinamica, le previsioni del VaR risultano essere sufficientemente accurate in quanto non è possibile rifiutare l'ipotesi nulla in nessuno dei test riportati. Gli stessi risultati vengono raggiunti anche dal modello DCC con termini di innovazione gaussiani. Nella Tabella 3.3.6 viene riportato il numero di volte in cui viene superato il VaR a differenti livelli di confidenza per ciascun modello analizzato. Utilizzando innovazioni distribuite secondo una t all'interno di un modello DCC, si ottengono delle previsioni del VaR troppo prudenti che possono condurre ad assumere dei rischi troppo elevati (Tabella 3.3.6).

Il VaR stimato in questi casi fa riferimento a livelli di confidenza pari al 90%, 95% e 99%, quindi essendo i casi definiti sfavorevoli (colonne 2-4 della Tabella 3.3.5) al di sotto della soglia considerata, si può affermare che, i modelli con copulae dinamici ed il modello DCC-G riescono a determinare un ammontare di capitale sufficiente a coprire le perdite che si sarebbero manifestate nel periodo successivo.

Le ultime tre colonne della Tabella 3.3.5 fanno riferimento al test di indipendenza delle previsioni. Come è possibile notare, il modello DCC in questo caso rifiuta l'ipotesi nulla a livello 1% mentre ciò non avviene per i modelli R-vine dinamici. Ciò indica che la scelta di un modello con copula dinamico per la struttura di dipendenza risulta essere estremamente utile per le previsioni del Value at Risk. Questo risultato è in contrapposizione a quello raggiunto da

Almeida, Czado, and Manner (2016) che evidenziavano come una struttura dinamica nei parametri della copula non aveva alcuna influenza sulle capacità previsive del Value at Risk. Infine, i modelli R-vine e D-vine con copulae statiche risultano essere totalmente inappropriati all'interno del contesto considerato in quanto ottengono le performance peggiori rispetto a tutti gli altri modelli considerati.

Visti i buoni risultati ottenuti da un modello classico DCC con innovazioni Gaussiane perché dovrebbero essere utilizzati dei modelli con copulae? La risposta è in una riduzione del capitale di rischio, infatti i valori dei quantili ottenuti con modelli con copulae dinamiche risultano essere più bassi rispetto al modello DCC e ciò evidenzia una maggiore accuratezza nelle previsioni. Questo è dovuto al fatto che la distribuzione Gaussianiana dei termini di innovazione non tiene conto dei possibili eventi in coda. Anche nel caso di modello DCC- t i quantili di quest'ultimo risultano sempre essere più elevati rispetto ai modelli con copulae. È quindi possibile concludere che il modello introdotto in questo lavoro riesce a ridurre i livelli di capitale di rischio raggiungendo comunque gli stessi risultati in termini di previsioni del VaR.

Abbiamo visto come la metodologia presentata in questo lavoro può essere utilizzata in sostituzione ai più tradizionali modelli DCC riuscendo ad ottenere, in questo contesto, delle performance migliori sia a livello in-sample sia a livello di bontà delle previsioni. Nel successivo capitolo viene analizzata l'utilità della procedura all'interno di problematiche inerenti la costruzione di portafogli diversificati.

Dai risultati ottenuti, risulta evidente che, il modello proposto in questo studio, riesce ad ottenere delle previsioni più accurate sull'ammontare di capitale necessario a coprire le perdite future rispetto agli altri modelli considerati. Per il portfolio manager ciò risulta essere estremamente rilevante in quanto permette meglio di fronteggiare quelle situazioni che possono generare perdite nel portafoglio finanziario.

Tabella 3.3.4: Confronto delle log-likelihood ad un passo

Modello	Pred. Log-likelihood
R-vine-D	3536.44
Tr. R-vine-D (6)	3537.78
R-vine-st	3262.81
D-vine model	3198.43
DCC - Gauss	3423.66
DCC - t	3492.51

Tabella 3.3.5: Backtesting per le valutazioni delle previsioni

Modello	<i>Unconditional Coverage</i>			<i>Conditional Coverage</i>		
	10%	5%	1%	10%	5%	1%
R-vine-D	0.651	0.551	0.451	0.493	0.336	0.121
Tr. R-vine-D (6)	0.449	0.584	0.528	0.499	0.385	0.217
R-vine-st	0.476	0.061	0.023	0.786	0.045	0
D-vine model	0.503	0.055	0.012	0.866	0.074	0
DCC - Gauss	0.803	0.555	0.512	0.866	0.454	0
DCC - t	0.816	0.523	0	0.487	0.413	0

I valori riportati in tabella rappresentano i diversi *p-value* dei test *Unconditional Coverage* di Kupiec (1995) e *Conditional Coverage* di Christoffersen (1998)

Tabella 3.3.6: Number of Exceedances

Modello	10%	5%	1%
R-vine-D	40	21	5
Tr. R-vine-D (6)	40	21	5
R-vine-st	38	19	5
D-vine model	36	19	4
DCC - Gauss	38	19	5
DCC - t	30	10	1

Capitolo 4

Ottimizzazione di portafoglio

Al fine di validare a livello pratico il modello proposto, in questo capitolo viene effettuata un'analisi in ottica di allocazione di portafoglio. Nel paragrafo 2.2 sono stati presentati alcuni lavori che hanno riguardato l'utilizzo delle copulae all'interno di tecniche di asset allocation.

In letteratura sono presenti diverse tecniche di ottimizzazione di portafoglio che appartengono alla ben nota letteratura della *modern portfolio theory* introdotta da Markowitz (1952), dove la base teorica è costituita dal trade-off rischio rendimento. In tali modelli l'obiettivo è la massimizzazione del rendimento, dato un determinato livello di rischio, oppure minimizzare il rischio dato un rendimento atteso prestabilito. Quando si parla di rischio ci si riferisce generalmente a delle variazioni del valore di mercato di uno o più strumenti finanziari. Spesso un aumento della variabilità temporale nelle serie dei rendimenti è causato da quegli episodi che possiedono una bassa probabilità di manifestarsi ma con effetti piuttosto rilevanti sull'intera distribuzione della serie. All'interno di teorie di asset allocation, tali episodi generano un aumento della rischiosità dell'intero portafoglio. Trattandosi di episodi riguardanti il comportamento congiunto di diverse serie finanziarie, situati nelle code della stessa distribuzione congiunta, l'utilizzo delle copulae all'interno di tecniche di ottimizzazione di portafoglio può risultare estremamente utile.

Quando si parla di rischio, all'interno di teorie di asset allocation, ci si riferisce generalmente a quegli episodi che possiedono una bassa probabilità di manifestarsi ma con effetti rilevanti sull'intero portafoglio. Trattandosi di episodi riguardanti il comportamento congiunto di diverse serie finanziarie, situati nelle code della stessa distribuzione congiunta, l'utilizzo delle copulae all'interno di tecniche di ottimizzazione di portafoglio può risultare estremamente utile.

Di seguito vengono illustrate alcune tecniche di ottimizzazione, nello specifico l'approccio media-varianza di Markowitz (1952), il *Global Minimum Variance*, la minimizzazione del *Value at Risk* (VaR) e del *Conditional Value at Risk* (CVaR) (o *Expected Shortfall*). Successivamente vengono analizzate le performance di un portafoglio costruito utilizzando il metodo introdotto nel Capitolo 3 riportando un confronto con altre metodologie.

4.0.1 Media-Varianza (Markowitz, 1952)

Nel suo lavoro pionieristico Markowitz (1952) proponeva di misurare il rischio di un'attività finanziaria attraverso la variazione dei rendimenti di tale attività intorno al suo valore medio. L'obiettivo di tale lavoro è quello di individuare il portafoglio che minimizza il rischio dato un determinato vincolo di rendimento attraverso la diversificazione tra diverse attività finanziarie. La soluzione del suo modello porta alla formazione della frontiera di portafoglio che si configura come un'iperbole avente concavità rivolta verso destra con un punto di minimo che rappresenta il portafoglio con varianza minima all'interno di un sistema cartesiano con il rendimento atteso di portafoglio (R_p) in ordinata ed il rischio atteso di portafoglio (σ_p) in ascissa. Tutti i portafogli efficienti si collocano nel tratto crescente di tale frontiera e si identificano in tutte quelle combinazioni rischio-rendimento che si trovano al di sopra del portafoglio con varianza minima e che garantiscono il maggior rendimento possibile per ogni livello di rischio. Secondo l'approccio proposto da Markowitz (1952), un

investitore diversifica il proprio portafoglio sulla base di una sua determinata funzione di utilità. Il portafoglio ottimo viene individuato nel punto di tangenza tra la più alta curva di utilità possibile e la suddetta frontiera (Figura 4.0.1). Analiticamente, il problema di Markowitz è quello della ricerca di un minimo sotto vincolo, quindi:

$$\begin{cases} \min & \sigma_p^2 = w' \Sigma w \\ \text{s.t.} & \mu_p = w' \mu \\ & \iota' w = 1 \end{cases} \quad (4.0.1.1)$$

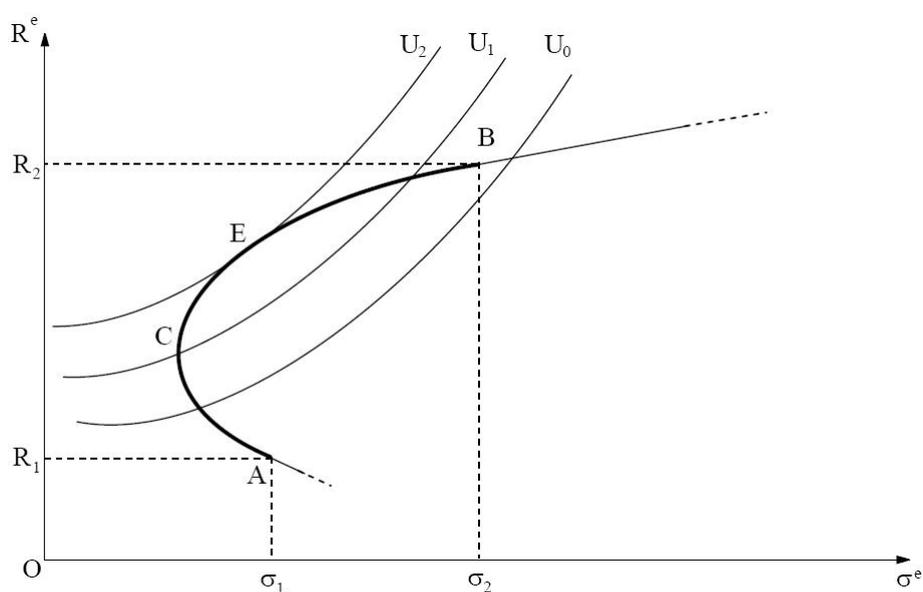
Il termine w rappresenta il vettore dei pesi, Σ la matrice varianze-covarianze di dimensione $n \times n$, μ il vettore dei rendimenti delle singole attività finanziarie di dimensione n , $\iota = [1, 1, \dots, 1]$ e n il numero di attività rischiose.

Tale modello è stato oggetto di critiche riguardanti l'utilizzo della matrice di covarianza, in quanto, data la sua natura simmetrica, essa considera in egual modo cambiamenti di segno positivo e di segno negativo ed anche l'utilizzo di funzioni di utilità CRRA e di distribuzione normale dei rendimenti.

4.0.2 Global Minimum Variance

A differenza dell'approccio media-varianza, la metodologia *Global Minimum Variance* si sofferma sulle previsioni della matrice delle covarianze, evitando di effettuare previsioni sul rendimento atteso di portafoglio.

Numerosi studi empirici hanno dimostrato come l'utilizzo di tecniche di allocazione di portafoglio basate sulle previsioni della media dei rendimenti, ottengono risultati inferiori rispetto a quelli ottenuti senza prevedere la stessa (Chan, Karceski, and Lakonishok, 1999). Tale metodologia risulta ampiamente utilizzata a livello empirico in associazione con modelli di previsione della matrice di covarianza come ad esempio il metodo DCC o il BEKK.

Figura 4.0.1: *Frontiera efficiente di portafoglio e punto di equilibrio*

NOTA: le curve di utilità (U_0, U_1, U_2) sono di tipo *Constant Relative Risk Aversion* (CRRA).

I punti A e B rappresentano due differenti portafogli. Il punto A corrisponde al portafoglio interamente investito nell'attività finanziaria con un rendimento pari a R_1 , mentre il punto B corrisponde al portafoglio interamente investito nell'attività finanziaria con un rendimento pari a R_2 .

I punti C e E sono portafogli ricavati attraverso la diversificazione. Il punto C rappresenta il portafoglio con varianza minima, mentre il punto E si riferisce al portafoglio di equilibrio (ottimo), ossia nel punto di tangenza della curva della frontiera di portafoglio alla più alta curva di utilità possibile.

Il problema di minimizzazione inerente a questa metodologia è il seguente:

$$\begin{aligned} \min_w \quad & \sigma_p = w' \Sigma w & (4.0.2.1) \\ \text{s.t.} \quad & \iota' w = 1 \end{aligned}$$

I pesi ottenuti dal metodo GMV vengono ricavati come

$$w = \frac{\Sigma^{-1} \mathbf{1}}{\mathbf{1}' \Sigma^{-1} \mathbf{1}}.$$

4.0.3 Media-Value at Risk (VaR)

La misura di rischio considerata negli approcci precedenti è la matrice di covarianza. Tale misura è stata oggetto di numerose critiche dovute alla sua natura simmetrica e questo ha spinto diversi investitori istituzionali alla creazione di nuove misure di rischio come ad esempio il Value at Risk (VaR). Il Value at Risk sintetizza la massima perdita attesa, su un orizzonte temporale prestabilito e nei limiti di un intervallo di confidenza:

$$Pr(\mu_p \leq -VaR_\alpha) = \alpha \quad (4.0.3.1)$$

dove μ_p è il rendimento di portafoglio, ed α è il livello di significatività. Da questa definizione emerge che maggiore è il VaR ad un determinato livello di confidenza, maggiore sarà il rischio di portafoglio.

Esistono diverse metodologie di calcolo del VaR come ad esempio il metodo media-varianza, il metodo delle simulazioni storiche e quello delle simulazioni Monte Carlo.

In questo approccio il problema di ottimizzazione diviene:

$$\begin{aligned} \min_w \quad & VaR_\alpha & (4.0.3.2) \\ \text{s.t.} \quad & w' \mu_i = \mu_0 \\ & \iota' w = 1 \end{aligned}$$

dove μ_0 rappresenta un livello di rendimento di portafoglio prefissato. Eliminando tale vincolo otteniamo un modello *min-VaR*.

Il Value at Risk, tutttavia presenta delle debolezze. La prima fra tutte è la non additività, proprietà indispensabile di ogni misura di rischio. Tale proprietà assicura che, attraverso l'aggregazione di due o più portafogli, la misura di rischio del portafoglio aggregato è minore rispetto alla somma delle singole misure di rischio. Nella metodologia VaR, invece, il VaR ottenuto attraverso la combinazione dei due portafogli può risultare superiore alla somma dei singoli VaR di portafoglio.

Il Value at Risk inoltre non fornisce alcuna informazione riguardante le perdite che eccedono la soglia prestabilita. Pertanto si è resa necessaria la creazione di una nuova misura di rischio il *Conditional Value at Risk* (CVaR).

4.0.4 Min-Conditional Value at Risk

Il *Conditional Value at Risk*, spesso indicato anche come *Expected Shortfall*, rappresenta la perdita attesa condizionata al fatto che si sia superata la soglia stabilita dal VaR. In sostanza essa rappresenta una media di tutte le possibili perdite superiori al livello di VaR (Figura 4.0.2)

$$CVaR_\alpha = E[X|X \geq VaR_\alpha] \quad (4.0.4.1)$$

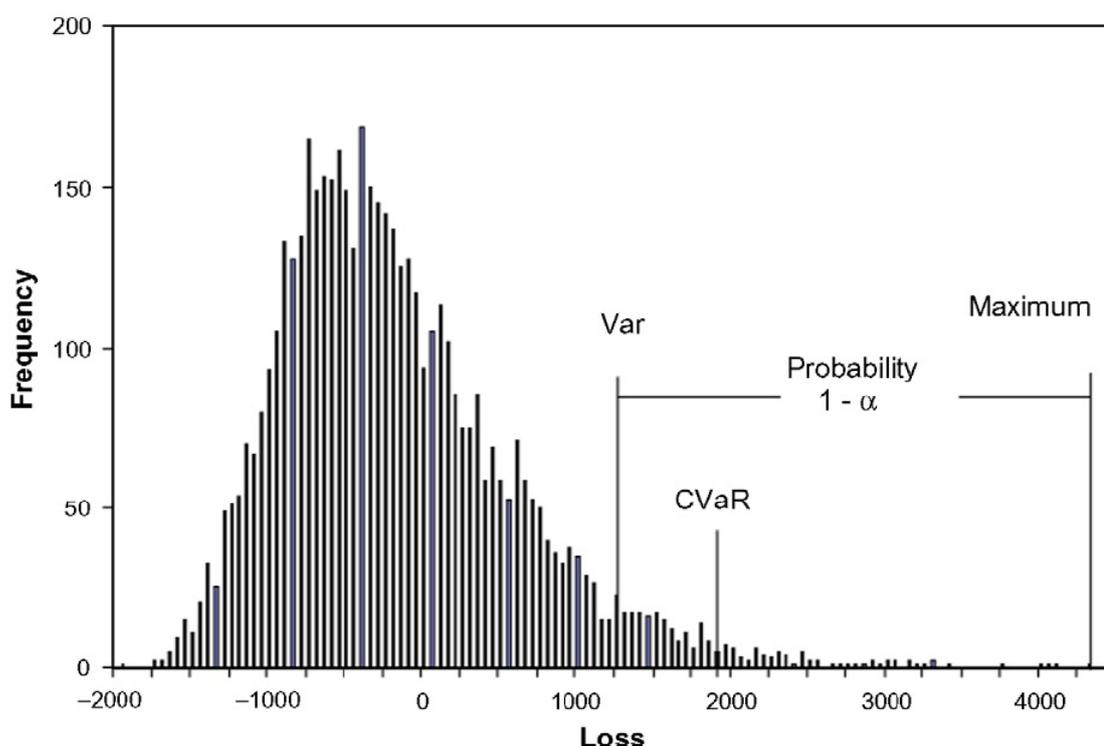
dove X rappresenta la distribuzione delle perdite attese.

Al contrario del VaR, il CVaR possiede la proprietà di sub-additività tipica delle misure di rischio.

Il problema di ottimo in questo caso consiste in:

$$\begin{aligned} \min_w \quad & CVaR_\alpha & (4.0.4.2) \\ \text{s.t.} \quad & w' \mu_i = \mu_0 \\ & \iota' w = 1 \end{aligned}$$

Figura 4.0.2: Rappresentazione VaR e CVaR



Fonte: Webby et al (2007)

4.1 Indici di performance di portafoglio

Le valutazioni dei risultati raggiunti da diversi portafogli si basano sostanzialmente su due elementi, il rendimento conseguito ed il livello di rischio di portafoglio. Gli indicatori basati su questi elementi vengono definiti *risk-adjusted*. Il più comune è rappresentato dallo *Sharpe Ratio* (SR), costituito dal rapporto tra il rendimento in eccesso di portafoglio rispetto ad un tasso *risk-free* e la deviazione standard di portafoglio:

$$SR = \frac{\mu_{ex}}{\sigma_p} \quad (4.1.0.1)$$

dove il rendimento in eccesso $\mu_{ex} = \mu_p - r_f$ e σ_p è la deviazione standard del portafoglio. Un valore maggiore dello *Sharpe Ratio* implica una migliore performance di portafoglio. Questo è un indice simmetrico ovvero considera

allo stesso modo scostamenti positivi e negativi dalla media.

Un altro indice di performance è costituito dall'indice di Treynor. Il calcolo di quest'ultimo è simile a quello dello Sharpe Ratio dove però viene utilizzato il β_p di portafoglio¹ in sostituzione alla deviazione standard.

$$Tr = \frac{\mu_{ex}}{\beta_p} \quad (4.1.0.2)$$

Tale indice presuppone che il rischio specifico associato a ciascun titolo sia totalmente eliminato da una buona diversificazione.

Un ulteriore indice costruito prendendo in considerazione il β_p di portafoglio è l'*alpha* di Jensen.

L'*alpha* rappresenta la differenza tra i rendimenti di un determinato portafoglio e i rendimenti del portafoglio di benchmark:

$$\alpha_p = (\mu_p - rf) - \beta_p(\mu_b - rf) \quad (4.1.0.3)$$

con μ_p , rf e μ_b che rappresentano rispettivamente il rendimento del portafoglio dell'investitore, il rendimento dell'attività priva di rischio ed il rendimento del portafoglio di benchmark.

Negli studi empirici l'utilizzo di tale indice ha ottenuto un notevole riscontro nonostante alcuni svantaggi legati all'uso della sola misura di rischio sistematico. Alexander and Baptista (2010) utilizzano tale misura nella costruzione di una frontiera efficiente *alpha-TEV*.

Uno dei maggiori problemi nell'utilizzare la deviazione standard come misura di rischio è che essa considera allo stesso modo le fluttuazioni positive e negative rispetto alla media. L'indice di Sortino, diversamente dall'indice di Sharpe, considera il *downside-risk* come misura di rischio. Per calcolare

¹Indice di rischio sistematico calcolato come rapporto tra la covarianza tra il rendimento di portafoglio ed il rendimento di mercato, e la varianza del rendimento di mercato. In formule:

$$\beta_p = \frac{\sigma_{pm}}{\sigma_m}$$

questo indice c'è bisogno di fissare un tasso di rendimento minimo accettabile (Minimal Acceptable rate of Return (MAR)). Generalmente questo tasso viene fissato ad un livello pari al tasso di rendimento risk free oppure viene posto a livello zero.

$$S_o = \frac{\mu_p - MAR}{DSR} \quad (4.1.0.4)$$

DSR indica il *Downside Risk* definito come:

$$DSR = \sqrt{\frac{1}{N} \sum_{i=1}^N (\min(0, r_i - MAR))^2}$$

dove r_i rappresenta il rendimento giornaliero del portafoglio. La funzione min assicura che, se si osservano dei rendimenti superiori al MAR , questi non vengono presi in considerazione nel calcolo del *Downside Risk*.

Tutti gli indici appena descritti verranno utilizzati all'interno di questo lavoro.

4.2 Portafoglio con copula R-vine dinamica

In questa sezione viene costruito un portafoglio diversificato utilizzando un copula R-vine dinamica all'interno della procedura min-VaR descritta nei paragrafi precedenti. Viene considerata una gestione di portafoglio attiva, pertanto il portafoglio viene ribilanciato con cadenza prestabilita.

L'obiettivo principale di una gestione attiva di portafoglio è quella di trovare l'allocazione di portafoglio "migliore" per i giorni futuri ovvero i pesi ottimali di portafoglio per ogni singolo asset.

Definendo μ_{t+1} il rendimento di portafoglio futuro e Σ_{t+1} la matrice di covarianza futura, il problema di ottimo in un modello media-varianza diviene:

$$\begin{aligned} \min_w \quad & w' \Sigma_{t+1} w \\ \text{s.t.} \quad & w' \mu_{t+1} = \mu_0 \\ & \iota' w = 1 \end{aligned}$$

Ovviamente in tale modello i rendimenti futuri e la matrice di covarianza futura non sono noti. In questo ambito viene inserita la procedura R-vine dinamica descritta nel capitolo 3, effettuando delle previsioni ad un passo all'interno di una *rolling window*. Nel metodo rolling la dimensione della finestra temporale è fissa e le informazioni vengono aggiornate mano a mano sottraendo le prime osservazioni ed aggiungendo le più recenti. In altre parole, mano a mano che la finestra viene spostata in avanti, si tralascia l'informazione più vecchia e si aggiunge l'informazione più recente.

La procedura viene effettuata nella porzione out-of-sample del dataset (375 osservazioni).

L'ampiezza della finestra è fissata pari a 1400 ed il ribilanciamento dei pesi viene effettuato settimanalmente. Pertanto l'aggiornamento informativo viene effettuato sottraendo 5 osservazioni (numero di giorni di contrattazione in borsa settimanale) dall'inizio del dataset ed aggiungendo lo stesso numero di osservazioni alla fine dello stesso, calcolando di volta in volta i pesi ottimali di portafoglio.

L'intera procedura è così composta:

1. Stima di un modello GARCH(1,1) per ogni singola serie storica con residui estratti da una distribuzione t di Student.
2. I residui standardizzati ottenuti nel punto precedente vengono trattati con una copula multivariata R-vine dinamica come descritto nel capitolo 3.
3. Le stime dei parametri dei modelli GARCH e della copula multivariata vengono utilizzate per effettuare delle previsioni ad un passo delle serie dei rendimenti. Vengono effettuate 10000 simulazioni di Montecarlo di modo da ottenere 10000 previsioni ad un passo per ogni singola serie.

4. Le previsioni vengono utilizzate per calcolare il rendimento atteso di portafoglio e la matrice di covarianza.
5. La riallocazione di portafoglio viene effettuata settimanalmente, quindi si ha bisogno di prevedere 5 rendimenti ogni volta, per poi muovere in avanti la finestra temporale (rolling windows)

Le previsioni vengono inserite all'interno della procedura min-VaR (4.0.3.2) ottenendo:

$$\begin{aligned} \min_w \quad & VaR_{\alpha,t+1} \\ & \iota'w = 1 \end{aligned}$$

È facile notare come la procedura appena descritta possa essere utilizzata all'interno di ciascuna delle tecniche di allocazione di portafoglio descritte nei paragrafi precedenti.

4.2.1 Risultati

In questo paragrafo vengono mostrati i risultati ottenuti all'interno del dataset considerato. Il portafoglio equipesato viene ricavato attraverso una strategia *naive* con pesi pari a $\frac{1}{m}$, dove m è il numero totale dei titoli azionari presi in considerazione. Per quanto riguarda la costruzione dei diversi indici *risk adjusted* quali lo Sharpe Ratio, l'indice di Sortino, l'indice di Treynor e dell'alpha di Jensen si considera un tasso privo di rischio, r_f , pari a zero, mentre il benchmark, r_m , è rappresentato dal portafoglio costruito con un modello tradizionale DCC con innovazioni gaussiane al fine di valutare la reale utilità dei modelli con copula dinamici.

I singoli portafogli, ad eccezione del portafoglio equipesato, sono costruiti con la metodologia min-VaR. Nella costruzione dei portafogli vengono effettuate diverse assunzioni:

- impossibilità di effettuare vendite allo scoperto;

- limite di concentrazione su ogni singolo titolo pari al 16% ;
- zero costi di transazione.

La percentuale riguardante il limite di concentrazione è stata determinata effettuando diverse prove di allocazione i cui risultati hanno evidenziato che, per percentuali maggiori del 16%, non vi è alcun incremento nelle performance di portafoglio. Nella Tabella 4.2.1 vengono mostrati i risultati ottenuti dai vari portafogli.

Dal confronto emergono diversi risultati. Innanzitutto è evidente come il portafoglio costruito con una metodologia *naive* ottenga dei rendimenti medi superiori rispetto agli altri portafogli considerati. I portafogli con copula dinamici, ottenuti attraverso un'ottimizzazione min-VaR, ottengono dei rendimenti medi di poco inferiori al modello equipesato ma raggiungono le migliore performance assolute in termini di deviazione standard di portafoglio e di VaR. Risulta chiaro il vantaggio apportato dai modelli con copulae dinamiche nello studio della dipendenza nei confronti dei più classici modelli DCC. Ciò conferma la bontà, anche in chiave economica, della metodologia introdotta in questo lavoro. Anche gli indici di performance *risk adjusted* confermano come tali modelli sono preferibili agli altri modelli con pesi ribilanciati. In questo contesto, invece, i modelli con copulae statiche non presentano alcuna utilità risultando essere inferiori anche ai modelli DCC.

Tabella 4.2.1: Confronto performance portafogli min-VaR

	$\mu_p\%$	$\sigma_p\%$	Treydor	SR	So	alpha	VaR		CVaR	
							95%	99%	95%	99%
Equal	0.207	1.0568	0.00335	0.00196	0.00184	0.00199	2.487	3.751	3.386	4.215
R-vine-D	0.1786	0.8216	-0.00223	0.00186	0.00187	0.00196	1.009	1.974	1.306	2.187
Tr. R-vine-D (6)	0.1805	0.7991	-0.00253	0.00184	0.00186	0.00195	0.992	1.791	1.243	2.204
R-vine-st	0.0816	0.9921	-0.0016	0.00105	0.00113	0.00127	2.238	3.066	2.955	3.538
D-vine	0.1179	0.9501	-0.00239	0.00117	0.00162	0.00161	2.124	2.945	2.902	3.373
DCC - Gauss	0.1305	0.9589	-0.00228	0.00109	0.00121	0.00169	1.811	2.686	2.528	3.184
DCC - t	0.1297	0.9605	-0.00281	0.00098	0.00125	0.00181	1.707	2.396	2.274	3.008

Capitolo 5

Conclusioni

L'obiettivo principale del lavoro proposto è quello di indagare l'utilizzo delle copulae multivariate all'interno di problematiche di asset allocation. In tale contesto viene sviluppato un modello con copula multivariata dinamica di tipo Regular-Vine per lo studio delle dinamiche di dipendenza esistenti tra un gruppo di 17 serie di rendimenti finanziari di titoli quotati all'interno dell'indice di mercato FTSE-mib.

La scelta di utilizzare la copula R-vine è dovuta alla sua maggiore flessibilità rispetto ai sottomodelli D-vine e C-vine. Essa infatti non vincola la struttura di dipendenza su un percorso prestabilito, permettendone così un'analisi più dettagliata ma anche più complessa. In letteratura esistono diverse tipologie di modelli che considerano delle dinamiche temporali all'interno dei modelli con copulae come ad esempio i modelli *regime-switching* ed i modelli *factor copula*. In letteratura però non risultano essere presenti lavori in cui viene analizzata una struttura di dipendenza dinamica di tipo R-vine. Il lavoro proposto si inserisce proprio in questo contesto, utilizzando una struttura dinamica nei parametri di tipo *Generalized Autoregressive Score*. Questo lavoro può essere visto come un'estensione del lavoro proposto da Almeida, Czado, and Manner (2016), in cui gli autori utilizzano una copula multivariata D-vine.

Nell'analisi svolta viene indagata l'utilità del modello applicato a problemi

di ottimizzazione di portafoglio. Il modello proposto viene comparato con altri modelli già ben noti in letteratura ed ampiamente utilizzati come il modello DCC e due modelli con copula multivariata statica (D-vine ed R-vine).

I risultati ottenuti mostrano che il modello con copula multivariata dinamica ottiene dei buoni risultati sia nella porzione in-sample del campione sia nelle previsioni ad un passo del Value at Risk. Ciò indica che l'introduzione di una dinamicità nella struttura di dipendenza è di grande rilevanza nello studio delle dinamiche congiunte esistenti tra diverse variabili finanziarie, diversamente da quanto raggiunto da Almeida, Czado, and Manner (2016). Per tale ragione il modello proposto risulta essere una valida alternativa al più noto modello DCC.

Dall'analisi di allocazione di portafoglio, condotta attraverso la metodologia min-VaR, è emerso che il modello R-vine dinamico ottiene le migliori performance a livello di rischiosità di portafoglio (deviazione standard di portafoglio, VaR e CVaR), essendo questa inferiore rispetto ad un portafoglio equipeso, un portafoglio con modelli DCC ed ai portafogli costruiti con copula multivariata statica. Il portafoglio equipeso ottiene, invece, dei livelli di rendimento medio superiori agli altri modelli considerati, anche se di poco. Nella scelta tra l'utilizzo di modelli classici come il DCC ed un modello con copula multivariata dinamica, c'è bisogno di effettuare un'analisi riguardante i costi ed i benefici. In letteratura è noto che l'utilizzo di modelli con copula multivariata comporta un aumento del numero di parametri da stimare con il conseguente aumento del tempo di stima. L'introduzione di una dinamicità nei parametri accentua ancora di più questo svantaggio. In questo lavoro vengono comparati i tempi macchina che un normale computer impiega per la stima dei vari modelli messi in competizione. Come già evidenziato, i modelli con copulae presentano un grande svantaggio rispetto ai modelli DCC, impiegando un quantitativo di tempo da 2 a 10 volte superiore a quest'ultimi. Al fine di ovviare a questo problema, viene indagata la possibilità di troncamento del modello, come proposto da Brechmann, Czado, and Aas (2012). I risultati

mostrano come sia possibile troncare il modello al sesto livello, impostando le successive copulae presenti nella struttura R-vine come indipendenti, riuscendo così a ridurre notevolmente la complessità dell'intera procedura e di conseguenza anche i tempi macchina di stima.

Il modello introdotto in questo studio è stato applicato all'interno di una procedura di ottimizzazione di portafoglio di tipo min-VaR. Vengono utilizzati i rendimenti giornalieri di 17 titoli azionari facenti parte dell'indice di mercato FTSE-mib e viene effettuata una riallocazione settimanale. Visti i buoni risultati ottenuti e visti i tempi macchina di stima dell'intera procedura, si può affermare che tale modello è facilmente applicabile all'interno di un sistema di riallocazione settimanale e mensile. Più difficoltosa risulta invece la riallocazione giornaliera soprattutto in presenza di un numero elevato di titoli finanziari.

Nel contesto analizzato emerge una possibile non linearità nella relazione tra la performance e la complessità del modello: il passaggio da un meno complesso DCC ad un più complesso modello con copulae statiche peggiora la performance del portafoglio, ma un ulteriore aumento della complessità tramite l'uso di copulae dinamiche inverte questo risultato. Tale non linearità è evidenziata anche nella procedura di troncamento dove una riduzione della complessità del modello conduce alle medesime, o anche migliori, performance di portafoglio del modello completo.

L'intera procedura di stima del modello è stata implementata all'interno del programma econometrico *gretl* ed una parte del codice utilizzato viene riportata in Appendice.

Sviluppi futuri possono riguardare:

- l'utilizzo di altre tipologie di copulae bivariate all'interno della procedura PCC. È possibile indagare la presenza di una struttura di dipendenza asimmetrica attraverso l'utilizzo, ad esempio, della copula *skew-t*, o combinazioni convesse di esse;

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

- l'impiego di modelli diversi dal GARCH nella trattazione delle serie univariate e l'utilizzo di altre distribuzioni dei residui (skewed-t, GED);
- l'applicazione di altre tecniche di allocazione di portafoglio come ad esempio il min-CVaR ed altri indici di performance;
- l'incremento del numero di titoli oggetto di analisi e lo studio del modello con l'introduzione di indici obbligazionari e/o indici su commodities.

Appendice

Codice gretl

Il codice mostrato di seguito effettua la stima delle copulae bivariate statiche e dinamiche utilizzate nel modello. La parte riguardante le copulae statiche è stata oggetto di presentazione alla *5th gretl Conference*. Il pacchetto ancora non è stato reso disponibile in quanto la stima della copula t risulta essere piuttosto lenta rispetto alle altre e pertanto si sta cercando di ottimizzare la procedura.

```
#INPUT: 2 time series
#OUTPUT: parameter(s) estimation for 7 static and 7 dynamic bivariate copula
```

```
#-----
```

```
# This function returns the sample size (T) when the time series f1 and f2
# have the same starting date and end date. Otherwise, it returns zero.
```

```
function matrix match (series x, series y)
```

```
    scalar first1=firstobs(x)
```

```
    scalar first2=firstobs(y)
```

```
    if first1<$t1
```

```
        first1=$t1
```

```
    endif
```

```
    if first2<$t1
```

```
        first2=$t1
```

```
    endif
```

```
    scalar last1=lastobs(x)
```

```
    scalar last2=lastobs(y)
```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```

    if last1>$t2
        last1=$t2
    endif
    if last2>$t2
        last2=$t2
    endif
    scalar T1=last1-first1+1
    scalar T2=last2-first2+1
    string start1=obslabel(first1)
    string start2=obslabel(first2)
    string end1=obslabel(last1)
    string end2=obslabel(last2)
    matrix output= zeros(1,6)
    if start1=start2 && end1=end2
        matrix output={T1,T2,first1,last1,first2,last2}
        /*
        print " "
        printf "series %d: %s\n", 1, varname(1)
        printf "series %d: %s\n", 2, varname(2)
        print " "
        printf "using observations %s:%s (T = %g)\n",start1,end1,T1
    else
        print "Time series do not match: "
        printf " 1st series ranges from %s to %s (T=%g)\n",start1,end1,T1
        printf " 2nd series ranges from %s to %s (T=%g)\n",start2,end2,T2
        */
    endif
    return output
end function

# Empirical Distribution Function-----
function matrix empirical_cdf(series x,series y)
    T=nobs(x)
    matrix id = seq(1,T)'
    matrix prob=id/(T+1)

```

```

matrix X = {x}~id
matrix Y = {y}~id
X=msortby(X,1)~prob
Y=msortby(Y,1)~prob

matrix out=msortby(X,2)[:,3]~msortby(Y,2)[:,3]

return out
end function

## Information Criteria-----

function matrix IC1(matrix ecd1,scalar LL, scalar npar)
matrix out = zeros(1,3)
out[1] = -2*(-LL)+2*npar #aic
out[2]= -2*(-LL)+npar*log(rows(ecd1)) #bic
out[3] = -2*(-LL)+2*npar*log(log(rows(ecd1))) #hqc
return out
end function

## Normal Copula-----

## Normal Copula pdf

function matrix NormalCopula_pdf(matrix U,scalar RHO)
if RHO <-1 || RHO >1
funcerr "Correlation parameter must be in (-1, 1)"
endif
matrix U = invcdf(z,U)
scalar crho = 1 - RHO^2
matrix w = sumr(U.^2)
matrix out1 = (1/sqrt(crho))*exp(-0.5/crho * (w - 2*RHO*U[:,1].*U[:,2])) .* exp(0.5*w)
return out1
end function

# Normal Copula cdf

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```
function matrix NormalCopula_cdf(matrix U, scalar RHO)
    if RHO <-1 || RHO >1
        funcerr "Correlation parameter must be in (-1, 1)"
    endif
    matrix U = invcdf(z,U)
    matrix out = cdf(D, RHO, U[,1], U[,2])
    return out
end function

## Normal Copula likelihood

function matrix NormalCopula_LL(matrix U, scalar RHO)
    matrix U = invcdf(z,U)
    scalar crho = 1 - RHO^2
    matrix w = sumr(U.^2)
    matrix LL = -1/(2*crho)*(w-2*RHO*U[,1].*U[,2])+0.5*w - (0.5*log(crho))
    return LL
end function

## Normal Copula estimation

function matrix NormalCopula_es(series x, series y, bool emp[0])
    if emp ==1
        matrix ecd1=empirical_cdf(x,y)
    else
        matrix ecd1={x,y}
    endif
    matrix s=zeros(1,9)
    RHO = corr(invCDF(z,ecd1[,1]),invCDF(z,ecd1[,2]))
    mle LLT = NormalCopula_LL(ecd1,RHO)
    params RHO
end mle --quiet
s[1] = RHO
s[9] =$stderr
s[2]=NA
```

```

    s[3]=-sum(NormalCopula_LL(ecd1,s[1]))
    s[4:6]=IC1(ecd1,s[3],1)
    return s
end function

## tStudent Copula-----

## t-Student Copula Likelihood

# arctan is used as reparametrization of correlation parameter
function matrix tCopula_LL(matrix U, scalar rho, scalar nu)#scalar nu)
    matrix U = invcdf(t,nu,U)
    scalar rho = tanh(rho)
    scalar crho = 1 - rho^2
    w = sumr(U.^2)
    m = (nu+2)/2
    k = (nu+1)/2
    matrix LL = log(gammafun(m))+log(gammafun(nu/2))-2*log(gammafun(k))-0.5*log(crho)
    LL = LL - m *(log(1+(w-2*rho*U[,1].*U[,2])/(nu*crho))) +
        + k*(log(1+U[,1].^2/nu)+log(1+U[,2].^2/nu))
    flush
    return LL
end function

## t-Student Copula pdf
function matrix tCopula_pdf(matrix U, scalar rho, scalar nu)
    if rho<-1 || rho>1
        funcerr "Correlation parameter must be in (-1, 1)"
    elif nu<2.1
        funcerr "Dof parameter must be in (2.1, 100)"
    endif
    if nu > 30
        print "For Dof greater than 100 Normal Copula pdf is used"
        out1 = NormalCopula_pdf(U,rho)
    else
        U = invcdf(t,nu,U)

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```

    scalar crho= 1 - rho^2
    w = sumr(U.^2)
    m = (nu+2)/2
    matrix out1 = gammafun(m)/gammafun(nu/2)/(nu*pi*sqrt(crho))*
                * ((1+(w-2*rho*U[,1].*U[,2])./(nu*crho)).^(-m))
    out1 = out1./(pdf(t,nu,U[,1]).*pdf(t,nu,U[,2]))
endif
return out1
end function

## t-Copula cdf

# Random matrices from the multivariate t distribution.
function matrix mvtrnd(matrix C, scalar df, scalar nob)
    T = cholesky(C)

    #generate normal and sqrt(normalized chi-square), then divide
    n = mnormal(nob, rows(T)) *T
    x = sqrt(mrandgen(c,df,nob,1)./df)
    matrix y = ones(nob,2).*x
    r = n./y
    return r
end function

# Numerically approximates the double integral of a bivariate Student's distribution,
# with zero means and unit variances, correlation rho and degrees of freedom df.
function matrix bivartcdf(matrix U, scalar rho, scalar df)
    nob=12500
    m = rows(U)
    matrix mrho = {1,rho;rho,1}
    matrix out = ones(m,1)
    loop i = 1..m
        r = mvtrnd(mrho,df,nob)
        r = r*sqrt((df-2)/df)
        matrix v = (mreverse(r'))'
        p = r | -r | v | -v
    end loop
end function

```

```

        out[i,1] = mean((p[,1].<=U[i,1]).*(p[,2].<=U[i,2]))
    endloop
    return out
end function

# t-Student Cumulative Distribution Function

function matrix tCopula_cdf(matrix U, scalar rho, scalar df)
    if rho<-1 || rho>1
        funcerr "Correlation parameter must be in (-1, 1)"
    elif df<2.1
        funcerr "Dof parameter must be in (2.1, 30)"
    endif
    if df > 30
        print "For df greater than 30 Normal Copula cdf is used"
        out1 = NormalCopula_cdf(U,rho)
    else

        U = invcdf(t,df,U)*sqrt((df-2)/df)
        out = bivartcdf(U,rho,df)
    endif
    return out
end function

## t-Student Copula estimation
function matrix tCopula_es(series x, series y, scalar emp)
    nu=10
    nn=nobs(x)
    if emp==1
        matrix ecd1=empirical_cdf(x,y)
    else
        matrix ecd1={x,y}
    endif
    series a = ecd1[,1]
    series b = ecd1[,2]
    scalar kt = ktau(a,b)

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```

scalar kt = atanh(kt)
catch mle LLT = check ? tCopula_LL(ecd1,kt,nu):NA
    scalar check = nu >2.1 && nu < 30
    params kt nu
end mle --quiet
err = $error
matrix s = zeros(1,8)

if err > 0 #Bound reached. Take the last value of the estimation procedure
    kt = tanh(kt)
    s[1]=kt
    s[2]=nu
    s[3] = -sum(tCopula_LL(ecd1,kt,nu))
    s[4:6] = IC1(ecd1,s[3],2)
    matrix ss = NA
    #print " "
    #print "Bound reached (gradient is not zero)"
else

    s[1] = tanh($coeff(kt))
    s[2] = $coeff(nu)
    ss = $stderr #deve essere corretto con il delta method
    co = $vcv
    s[3] = -$lnl
    s[4:6] = {$aic, $bic, $hqc}
endif

tail= -sqrt((s[2]+1)*(1-s[1])/(1+s[2]))
s[7:8] = 2*(1-cdf(t,s[2]+1,abs(tail)))
res = s~ss'
return res
end function

## Clayton Copula-----

```

```

#Log-Likelihood
function matrix CltCopula_LL(matrix U, scalar theta)
    if theta < 0.001
        theta = 0.001
    endif
    powu = U.^(-theta)
    lnu = log(U)
    logC = (-1/theta)*ln(sumr(powu)-1)
    logy = ln(theta+1) + (2*theta+1).*logC - (theta+1).*sumr(lnu)
    LL = logy
    return LL
end function

#Clayton Density Function
function matrix Clt_pdf(matrix U, scalar theta)
    if theta < 0
        funcerr "Clayton Copula's parameter must be positive"
    endif
    eta = 1+theta
    matrix cl_pdf= eta.*(U[,1].*U[,2]).^(-eta).*(sumr(U.^(-theta))-1).^(-(2+1/theta))
    return cl_pdf
end function

#Clayton Cumulative Distribution Function
function matrix Clt_cdf(matrix U, scalar theta)
    # at theta = 0 Clayton Copula is defined as the "independence copula"
    if theta < 0
        funcerr "Clayton Copula's parameter must be positive"
    endif
    if theta == 0
        matrix cl_cdf = U[,1].*U[,2]
    else
        matrix cl_cdf = ones(rows(U),1)
        loop i = 1..rows(U)
            cl = (U[i,1]^(-theta) + U[i,2]^(-theta)-1)^(-1/theta)
            cl_cdf[i,1] = xmax(cl,0)
        endloop
    end
end function

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```

        endloop
    endif
    return cl_cdf
end function

#Clayton Estimation
function matrix ClaytonCopula_es(series x,series y, bool emp[0])
    if emp==1
        matrix ecd1=empirical_cdf(x,y)
    else
        matrix ecd1={x,y}
    endif
    theta = 0.001
    mle LLCt = check ? CltCopula_LL(ecd1,theta):NA
        scalar check = theta > 0
        params theta
    end mle --quiet
    matrix s = zeros(1,9)
    s[1]=(theta)
    s[2] = NA
    s[3] = -sum(CltCopula_LL(ecd1,s[1]))
    s[4:6] = IC1(ecd1,s[3],1)
    s[9] = $stderr
    s[7] = 2^(-1/s[1])
    return s
end function

## Rotated Clayton Copula-----

function matrix RClaytonCopula_es(series x, series y, bool emp[0])
    if emp==1
        matrix ecd1=empirical_cdf(x,y)
    elif emp==0
        matrix ecd1={x,y}
    endif
    theta = 0.001

```

```

mle LLRCt = check ? CltCopula_LL(1-ecd1, theta):NA
    scalar check = theta > 0
    params theta
end mle --quiet
matrix s = zeros(1,9)
s[1]=theta
s[2] = NA
s[3] = -sum(CltCopula_LL(1-ecd1,s[1]))
s[4:6]=IC1(ecd1,s[3],1)
s[9] = $stderr
s[8] = 2^(-1/s[1])
return s
end function

##Frank Copula-----

#Log-likelihood
function matrix FrankCopula_LL(matrix U, scalar theta)
    expu = exp(theta.*U)
    if abs(theta) < 1e-5
        matrix LL = 2*theta*(prodr(U-.5))
    else
        matrix LL = log(-theta.*(exp(-theta)-1)) + theta.*sumr(U)
        LL = LL -2*log(abs(1+exp(theta.*(sumr(U)-1))-sumr(expu)))
    endif
    return LL
end function

#Density function
function matrix frank_pdf(matrix U, scalar theta)
    if theta == 0
        funcerr "Frank Copula's parameter cannot be zero"
    endif
    matrix fr_pdf = (theta*(1-exp(-theta)).*exp(-theta.*(U[,1]+U[,2])))./((1-exp(-theta))-
        - (1-exp(-theta.*U[,1])).*(1-exp(-theta.*U[,2]))).^2
    return fr_pdf

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```
end function

#Cumulative Distribution Function
function matrix frank_cdf(matrix U, scalar theta)
    if theta == 0
        funcerr "Frank Copula's parameter cannot be zero"
    endif
    eta = exp(-theta)
    matrix fr_cdf = -(1/theta)*log(1+(((eta.^U[,1])-1).*((eta.^U[,2])-1))./(eta-1))
    return fr_cdf
end function

#Frank Copula Estimation
function matrix FrankCopula_es(series x, series y, bool emp[0])
    if emp==1
        matrix ecd1=empirical_cdf(x,y)
    else
        matrix ecd1={x,y}
    endif
    scalar theta = 1
    catch mle LLFR = FrankCopula_LL(ecd1,theta)
        #scalar check = theta <> 0
        params theta
    end mle --quiet
    matrix s = zeros(1,9)
    s[1] = $coeff
    s[9] = $stderr
    s[2] = NA
    s[3] = -$lnl
    s[4:6] = {$aic, $bic, $hqc}
    return s
end function

## Gumbel Copula-----

#Density function
```

```

function matrix gumbel_pdf(matrix U, scalar theta)
    if theta <= 1
        funcerr "Gumbel Copula's parameter has to be grater than 1"
    endif
    matrix v = -log(U[,1])
    matrix w = -log(U[,2])
    matrix W = -log(U)
    matrix V = sumr(W.^theta)
    gum_pdf = gumbel_cdf(U, theta).*(1/(U[,1].*U[,2])).*((v.*w).^theta-1)
    gum_pdf = gum_pdf.*((V)^(1/theta-2))
    gum_pdf = gum_pdf.*((V)^(1/theta) + theta - 1)
    return gum_pdf
end function

#Cumulative Distribution Function
function matrix gumbel_cdf(matrix U, scalar theta)
    if theta <= 1
        funcerr "Gumbel Copula's parameter has to be grater than 1"
    endif
    matrix V = -log(U)
    matrix Z = sumr(V.^theta)
    matrix gcdf = exp(-Z^(1/theta))
    return gcdf
end function

#Gumbel copula log-likelihood
function matrix gCopula_LL(matrix U, scalar theta)
    theta = exp(theta)+1
    v = -log(abs(U))
    matrix vv = zeros(rows(v),2)
    loop i = 1..rows(v) --quiet
        vv[i,] = sort(v[i,])
    endloop
    vmin = vv[,1]
    vmax = vv[,2]
    logvv = log(vv)

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```

nlogC = (vmax.*(1+(vmin./vmax).^theta).^(1./theta))
lognlogC = log(nlogC)
logy = log(theta - 1+nlogC) - nlogC + sumr((theta-1).*logvv + v) + (1-2*theta).*lognlogC
LL = logy
return LL
end function

```

#Gumbel mle Estimation

```

function matrix GumbelCopula_es(series x, series y, bool emp[0])
    if emp==1
        matrix ecd1=empirical_cdf(x,y)
    else
        matrix ecd1={x,y}
    endif
    theta = 1.1
    theta = ln(theta-1)
    mle LLCggg = gCopula_LL(ecd1,theta)
    params theta
    end mle --quiet
    matrix s = zeros(1,9)
    s[1] = exp($coeff)+1
    s[9] = $stderr
    s[2] = NA
    s[3] = -$lnl
    s[4:6] = {$aic, $bic, $hqc}
    s[8]=2-2^(1/s[1])
end function

```

Rotated Gumbel Copula

```

function matrix RGumbelCopula_es(series x, series y, bool emp[0])
    if emp==1
        matrix ecd1=empirical_cdf(x,y)
    else
        matrix ecd1={x,y}
    endif

```

```

theta = 1.1
theta = ln(theta-1)
mle LLCrg = gCopula_LL(1-ecd1, theta)
    params theta
end mle --quiet
matrix s = zeros(1,8)
s[1] = exp($coeff)+1
s[2] = NA
s[3] = -$lnl
s[4:6] = {$aic, $bic, $hqc}
s[7]=2-2^(1/s[1])
return s
end function

function matrix restrpar(scalar omega, scalar A, scalar B)
    matrix parms = zeros(3,1)
    parms[1,1] = omega
    parms[2,1] = exp(A)
    parms[3,1] = exp(B)/(1+exp(B))
    return parms
end function

## Dynamic Gaussian Copula
function matrix gas_norm (matrix *parms, matrix *U)
    matrix LL = zeros(rows(U),1)
    matrix vSt = zeros(rows(U),1)
    matrix vrho = zeros(rows(U),1)
    omega = parms[1,1]
    A = parms[2,1]
    B = parms[3,1]
    parms = restrpar(omega,A,B)
    omega = parms[1,1]
    A = parms[2,1]
    B = parms[3,1]
    vrho[1,1] = omega
    rhor = vrho[1,1]

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```

omega = omega*(1-B)
rho = (1-exp(-vrho[1,1]))/(1+exp(-vrho[1,1]))
x = U[1,1]^2+U[1,2]^2
y = U[1,1]*U[1,2]
LL[1,1] = - 0.5 *(log(1-rho^2)) - 0.5 *(rho^2*x - 2*rho*y)/(1-rho^2)
vSt[1,1] = (2/(1-rho^2))*(y-rho-rho*(x-2)./(1+rho^2))
vrho[1,1] = omega + A*vSt[1,1] + B * rhor
loop i = 2..rows(U) --quiet
    x = U[i,1]^2+U[i,2]^2
    y = U[i,1]*U[i,2]
    rho = (1-exp(-vrho[i-1,1]))/(1+exp(-vrho[i-1,1]))

    #Log likelihood Gaussian Copula

    LL[i,1] = - 0.5 *(log(1-rho^2)) - 0.5 *(rho^2*x - 2*rho*y)./(1-rho^2)
    #scaled score function
    vSt[i,1] = (2/(1-rho^2))*(y-rho-rho*(x-2)./(1+rho^2))

    #GAS recursion
    vrho[i,1] = omega + A*vSt[i,1] + B * vrho[i-1,1]
    #vrho[i,1] = atanh(vrho[i,1])
endloop --quiet
return LL#-LL[rows(U),1]/rows(U)
end function

function matrix est_norm_gas(series x,series y, bool emp[0])
    if emp==1
        matrix U = empirical_cdf(x,y)
    else
        matrix U = {x,y}
    endif
    U = invcdf(z,U)
    matrix parms = {0.003; -3; 5}
    mle plutoo = gas_norm(&parms, &U)
    params parms

```

```

end mle --quiet --lbfgs
omega = parms[1]
A = parms[2]
B = parms[3]
parmsr = restrpar(omega, A, B)
gas = gas_norm2(parmsr,U)
matrix s = zeros(1,9)
s[1] = (gas[2])
s[2] = NA
s[3] = -gas[1]
s[4:6] = {$aic, $bic, $hqc}
return s
end function

function matrix gas_norm2 (matrix parms, matrix U)
matrix LL = zeros(rows(U),1)
matrix vSt = zeros(rows(U),1)
matrix vrho = zeros(rows(U),1)
omega = parms[1,1]
A = parms[2,1]
B = parms[3,1]
vrho[1,1] = omega
rhor = vrho[1,1]
omega = omega*(1-B)
rho = (1-exp(-vrho[1,1]))/(1+exp(-vrho[1,1]))
x = U[1,1]^2+U[1,2]^2
y = U[1,1]*U[1,2]
LL[1,1] = - 0.5 *(log(1-rho^2)) - 0.5 *(rho^2*x - 2*rho*y)/(1-rho^2)
vSt[1,1] = (2/(1-rho^2))*(y-rho-rho*(x-2)./(1+rho^2))
vrho[1,1] = omega + A*vSt[1,1] + B * rhor
loop i = 2..rows(U) --quiet
x = U[i,1]^2+U[i,2]^2
y = U[i,1]*U[i,2]
rho = (1-exp(-vrho[i-1,1]))/(1+exp(-vrho[i-1,1]))
LL[i,1] = - 0.5 *(log(1-rho^2)) - 0.5 *(rho^2*x - 2*rho*y)./(1-rho^2)
vSt[i,1] = (2/(1-rho^2))*(y-rho-rho*(x-2)./(1+rho^2))

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```

        vrho[i,1] = omega + A*vSt[i,1] + B * vrho[i-1,1]
    endloop --quiet
    return sumc(LL)~rho
end function

## Dynamic t-Student Copula
function matrix ct_gas (matrix parms, matrix ecd)
    nu = parms[4]
    UU = ecd
    matrix LL = zeros(rows(UU),1)
    matrix ScF = zeros(rows(UU),1)
    matrix vSt = zeros(rows(UU),1)
    matrix vrho = zeros(rows(UU),1)
    UU = invcdf(t,nu,ecd[1,])
    U = UU
    omega = parms[1,1]
    A = parms[2,1]
    B = parms[3,1]
    parms = restrpar(omega,A,B)
    omega = tanh(parms[1,1])
    A = parms[2,1]
    B = parms[3,1]
    vrho[1,1] = omega
    rhor = vrho[1,1]
    omega = omega*(1-B)
    scalar rho = tanh(vrho[1,1])
    scalar crho = 1 - rho^2
    x = U[1,1]^2 + U[1,2]^2
    y = U[1,1]*U[1,2]
    m = (nu+2)/2
    k = (nu+1)/2
    M = nu*crho+x-2*rho*y
    LL[1,1] = log(2)+(k*log(crho))-2*lngamma(k)+2*lngamma(nu/2)-((nu-2)/2)*log(nu)
    LL[1,1] = LL[1,1]+k*(log(nu+U[1,1]^2)+log(nu+U[1,2]^2))-m*log(M)
    ScF = -(nu+1)*(rho/crho)+(nu+2)*((nu*rho+y)/M)*(1/(1-rho^2))
    IM = -(nu+1)*(1+rho^2)/crho^2 + ((nu+2)*nu)/M + 2*(nu+2)*(nu*rho+y)^2/M*(1/(1-

```

```

vSt[1,1] = ScF*IM
vrho[1,1] = omega + A*vSt[1,1] + B * rhor
loop i = 2..rows(ecd) --quiet
    scalar crho = 1 - rho^2
    U = invcdf(t,nu,ecd[i,])
    x = U[1,1]^2+U[1,2]^2
    y = U[1,1]*U[1,2]
    m = (nu+2)/2
    k = (nu+1)/2
    M = nu*crho+x-2*rho*y
    LL[i,1] = log(2)+(k*log(crho))-2*lngamma(k)+2*lngamma(nu/2)-((nu-2)/2)*log(nu)
    LL[i,1] = LL[i,1]+k*(log(nu+U[1,1]^2)+log(nu+U[1,2]^2))-m*log(M)
    LL[i,1] = log(2)+(k*log(crho))-2*lngamma(k)+2*lngamma(nu/2)-((nu-2)/2)*log(nu)
    LL[i,1] = LL[i,1]+k*(log(nu+U[1,1]^2)+log(nu+U[1,2]^2))-m*log(M)
    IM = -(nu+1)*(1+rho^2)/crho^2 + ((nu+2)*nu)/M + 2*(nu+2)*(nu*rho+y)^2/M*(1/
    vSt[i,1] = ScF*IM
    vrho[i,1] = omega + A*vSt[i,1] + B * tanh(vrho[i-1,1])
endloop --quiet
return LL
end function

function matrix est_t_gas(series x, series y, scalar emp)
    if emp = 1
        U = empirical_cdf(x,y)
    else
        U = {x,y}
    endif
    ecd = U
    matrix parms = {0.5; -15; 5;10}
    parms[1] = atanh(parms[1])
    catch mle plutoo = check ? ct_gas(parms, ecd):NA
        scalar check = parms[4] > 2.1 && parms[4] < 50
        parms parms
    end mle
    omega = tanh(parms[1])
    A = parms[2]

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```

B = parms[3]
nu = parms[4]
parmsrr = restrpar(omega, A, B)
scalar omega = parmsrr[1]
scalar A = parmsrr[2]
scalar B = parmsrr[3]
matrix parmsr = {omega;A;B;nu}
gas = gas_t(parmsr,ecd)
matrix s = zeros(1,10)
s[1] = (gas[2])
s[2] = nu
s[3] = -gas[1]
s[4:6] = IC1(ecd,s[3],5)
return s
end function

function matrix gas_t(matrix parms, matrix ecd)
nu = parms[4]
UU = ecd
matrix LL = zeros(rows(UU),1)
matrix ScF = zeros(rows(UU),1)
matrix vSt = zeros(rows(UU),1)
matrix vrho = zeros(rows(UU),1)
UU = invcdf(t,nu,ecd[1,])
U = UU
omega = tanh(parms[1,1])
A = parms[2,1]
B = parms[3,1]
vrho[1,1] = omega
rhor = vrho[1,1]
omega = omega*(1-B)
scalar rho = vrho[1,1]
scalar crho = 1 - rho^2
x = U[1,1]^2 + U[1,2]^2
y = U[1,1]*U[1,2]
m = (nu+2)/2

```

```

k = (nu+1)/2
M = nu*crho+x-2*rho*y
LL[1,1] = log(2)+(k*log(crho))-2*lngamma(k)+2*lngamma(nu/2)-((nu-2)/2)*log(nu)
LL[1,1] = LL[1,1]+k*(log(nu+U[1,1]^2)+log(nu+U[1,2]^2))-m*log(M)
ScF = (-(nu+1)*(rho/crho)+(nu+2)*((nu*rho+y)/M))*(1/cosh(rho))^2
IM = (-(nu+1)*(1+rho^2)/crho^2 + ((nu+2)*nu)/M +
vSt[1,1] = ScF*IM
vrho[1,1] = omega + A*vSt[1,1] + B * rhor
loop i = 2..rows(ecd) --quiet
    scalar rho = tanh(vrho[i-1,1])
    scalar crho = 1 - rho^2
    U = invcdf(t,nu,ecd[i,])
    x = U[1,1]^2+U[1,2]^2
    y = U[1,1]*U[1,2]
    m = (nu+2)/2
    k = (nu+1)/2
    M = nu*crho+x-2*rho*y
    LL[i,1] = log(2)+(k*log(crho))-2*lngamma(k)+2*lngamma(nu/2)-((nu-2)/2)*log(nu)
    LL[i,1] = LL[i,1]+k*(log(nu+U[1,1]^2)+log(nu+U[1,2]^2))-m*log(M)
    ScF = (-(nu+1)*(rho/crho)+(nu+2)*((nu*rho+y)/M))*(1/cosh(rho))^2
    IM = (-(nu+1)*(1+rho^2)/crho^2 + ((nu+2)*nu)/M
    vSt[i,1] = ScF*IM
    vrho[i,1] = omega + A*vSt[i,1] + B * tanh(vrho[i-1,1])
endloop --quiet
return sumc(LL)
end function

# Dynamic Clayton Copula-----

function matrix der_cl(matrix theta, matrix U)
A = 1./U[,1].^theta + 1/U[,2].^theta - 1
lnU = ln(U)
UU = prodr(U)
Uth = 1./(U.^theta)
matrix ScF = (ln(A)./(theta.^2))+((-2-1./theta)).*((-Uth[,1].*lnU[,1]-Uth[,2].*lnU[,2])./A)
ScF = ln(UU)-ScF

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```

    ScF = Cl_pd(theta,U).*(1./(1+theta)-ScF)
    return ScF
end function

function matrix derLL_cl(matrix theta, matrix U)
    matrix der = der_cl(theta,U)
    der = der./Clt_pdf(theta,U)
    return der
end function

function matrix cClt_gas(scalar D, scalar E, scalar F, matrix UU)
    parms = {D;E;F}
    matrix U = UU[1,]
    matrix LL = zeros(rows(UU),1)
    matrix vSt = zeros(rows(UU),1)
    matrix thetag = zeros(rows(UU),1)
    omega = exp(parms[1,1])
    A = parms[2,1]
    B = parms[3,1]
    parms = restrpar(omega,A,B)
    omega = parms[1,1]
    A = parms[2,1]
    B = parms[3,1]
    thetag[1,1] = omega
    thetat = thetag[1,1]
    omega = omega*(1-B)
    matrix theta = ones(1,1).*thetat[1,1]
    powu = U.^(-theta)
    lnu = log(U)
    logC = (-1/theta)*ln(sumr(powu)-1)
    logy = ln(theta+1) + (2*theta+1).*logC - (theta+1).*sumr(lnu)
    LL[1,1] = logy
    matrix ScF = ones(rows(UU),1)
    ScF[1,1] = derLL_cl(theta,U)
    matrix IM = ones(rows(UU),1)
    IM[1,1] = fdjac(theta,derLL_cl(theta,U))
end function

```

```

vSt[1,1] = ScF[1,1]*1/sqrt(abs(IM[1,1]))
thetag[1,1] = omega + A*vSt[1,1] + B * thetat
loop i = 2..rows(UU) --quiet
    U = {UU[i,1],UU[i,2]}
    powu = U.^(-theta)
    lnu = log(U)
    logC = (-1/theta)*ln(powu[1,1]+powu[1,2]-1)
    logy = ln(theta+1) + (2*theta+1).*logC - (theta+1).*sumr(lnu)
    LL[i,1] = logy
    ScF[i,1] =derLL_cl(theta,U)
    IM[i,1] = fdjac(theta,derLL_cl(theta,U))
    vSt[i,1] = ScF[i,1]*1/sqrt(abs(IM[i,1]))
    thetag[i,1] = omega + A*vSt[i,1] + B * thetag[i-1,1]
endloop --quiet
return LL
end function

```

```

function matrix cClT_gas_par (matrix parms, matrix UU)
    matrix U = {UU[1,1],UU[1,2]}
    matrix LL = zeros(rows(UU),1)
    matrix vSt = zeros(rows(UU),1)
    matrix thetag = zeros(rows(UU),1)
    omega = parms[1,1]
    A = parms[2,1]
    B = parms[3,1]
    thetag[1,1] = omega
    thetat = thetag[1,1]
    omega = omega*(1-B)
    matrix theta = ones(1,1).*thetag[1,1]
    powu = U.^(-theta)
    lnu = log(U)
    logC = (-1/theta)*ln(sumr(powu)-1)
    logy = ln(theta+1) + (2*theta+1).*logC - (theta+1).*sumr(lnu)
    LL[1,1] = logy
    matrix ScF = ones(rows(UU),1)
    ScF[1,1] = derLL_cl(theta,U)

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```

matrix IM = ones(rows(UU),1)
IM[1,1] = fdjac(theta,derLL_cl(theta,U))
vSt[1,1] = ScF[1,1]*1/sqrt(abs(IM[1,1]))
thetag[1,1] = omega + A*vSt[1,1] + B * thetat
loop i = 2..rows(UU) --quiet
    U = {UU[i,1],UU[i,2]}
    theta = thetag[i-1,1]
    powu = U.^(-theta)
    lnu = log(U)
    logC = (-1/theta)*ln(powu[1,1]+powu[1,2]-1)
    logy = ln(theta+1) + (2*theta+1).*logC - (theta+1).*sumr(lnu)
    LL[i,1] = logy
    ScF[i,1] =derLL_cl(theta,U)
    IM[i,1] = fdjac(theta,derLL_cl(theta,U))
    #scaled score function
    vSt[i,1] = ScF[i,1]*1/sqrt(abs(IM[i,1]))
    thetag[i,1] = omega + A*vSt[i,1] + B * thetag[i-1,1]
endloop --quiet
return sumc(-LL)~theta
end function

function matrix est_clt_gas(series x, series y, bool emp[0])
    if emp==1
        matrix U = empirical_cdf(x,y)
    else
        matrix U = {x,y}
    endif
    matrix parms = {0.8; -3; 5}
    parms[1,1] = log(parms[1,1])
    catch mle plutoo = cClT_gas(parms, U)
        parms parms
    end mle --quiet --lbfgs
    omega = exp(parms[1])
    A = parms[2]
    B = parms[3]
    parmsr = restrpar(omega, A, B)

```

```

err = $error
matrix s = zeros(1,9)
gas = cClt_gas_par(paramsr,U)
s[1] = gas[2]
s[2] = NA
s[3] = gas[1]
s[4:6] = IC1(U,s[3],3)
return s
end function

# Dynamic Rotated Clayton

function matrix est_rclt_gas(series x, series y, bool emp[0])
  if emp==1
    matrix U = empirical_cdf(x,y)
  else
    matrix U = {x,y}
  endif
  matrix params = {0.8; -3; 5}
  params[1,1] = log(params[1,1])
  U = 1-U
  catch mle plutoo = cClt_gas(params, U)
    params params
  end mle --quiet --lbfgs
  omega = exp(params[1])
  A = params[2]
  B = params[3]
  paramsr = restrpar(omega, A, B)
  err = $error
  matrix s = zeros(1,9)
  gas = cClt_gas_par(paramsr,U)
  s[1] = gas[2]
  s[2] = NA
  s[3] = gas[1]
  s[4:6] = IC1(U,s[3],3)
  return s
end function

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```
end function
```

```
# Dynamic Frank Copula
```

```
function matrix der_fr(matrix theta, matrix U)
```

```
    thU = -theta.*U
```

```
    dt = U.*exp(thU)
```

```
    t = (1-exp(thU))
```

```
    expu = 1/(exp(theta))
```

```
    dercop = frank_pd(theta,U).*(1./theta+expu.*(1-expu).^(-1)-sumr(U)-2.*((1-expu)-prodr(t)).^
```

```
    fr = frank_pd(theta,U)
```

```
    dercop = -fr.^(-2).*dercop
```

```
    return dercop
```

```
end function
```

```
function matrix der_LLfr(matrix theta, matrix U)
```

```
    der_LL = der_fr(theta, U)./frank_pdf(U, theta)
```

```
    return der_LL
```

```
end function
```

```
function matrix cfrank_gas (matrix *parms, matrix *UU)
```

```
    matrix U = {UU[1,1],UU[1,2]}
```

```
    matrix LL = zeros(rows(UU),1)
```

```
    matrix vSt = zeros(rows(UU),1)
```

```
    matrix thetag = zeros(rows(UU),1)
```

```
    omega = parms[1,1]
```

```
    A = parms[2,1]
```

```
    B = parms[3,1]
```

```
    parms = restrpar(omega,A,B)
```

```
    omega = parms[1,1]
```

```
    A = parms[2,1]
```

```
    B = parms[3,1]
```

```
    thetag[1,1] = omega
```

```
    thetat = thetag[1,1]
```

```
    omega = omega*(1-B)
```

```

theta = (thetag[1,1])
expu = exp(theta.*U)
if abs(theta) < 1e-5
    LL[1,1] = 2*theta*(prodr(U-.5))
else
    LL[1,1] = log(-theta.*(exp(-theta)-1)) + theta.*sumr(U)
    LL[1,1] = LL[1,1] -2*log(abs(1+exp(theta.*(sumr(U)-1))-sumr(expu)))
endif
q = exp(-theta)
t1 = (1-exp(-theta*U[1,1]))
t2 = (1-exp(-theta*U[1,2]))
der_t1 = U[1,1]*exp(-theta*U[1,1])
der_t2 = U[1,2]*exp(-theta*U[1,2])
matrix ScF = ones(rows(UU),1)
ScF = (1/theta+q/(1-q)-(U[1,1]+U[1,2]))-2/((1-q)-t1*t2)*(q-der_t1*t1-der_t2*t2)
matrix IM = ones(rows(UU),1)
IM[1,1] = fdjac(theta,der_LLfr(theta, U))
vSt[1,1] = ScF*1/sqrt(IM)
thetag[1,1] = omega + A*vSt[1,1] + B * thetat
loop i = 2..rows(UU) --quiet
    theta = thetag[i-1,1]
    U = {UU[i,1],UU[i,2]}
    expu = exp(theta.*U)
    if abs(theta) < 1e-5
        LL[i,1] = 2*theta*(prodr(U-.5))
    else
        LL[i,1] = log(-theta.*(exp(-theta)-1)) + theta.*sumr(U)
        LL[i,1] = LL[i,1] -2*log(abs(1+exp(theta.*(sumr(U)-1))-sumr(expu)))
    endif
    q = exp(-theta)
    t1 = (1-exp(-theta*U[1,1]))
    t2 = (1-exp(-theta*U[1,2]))
    der_t1 = U[1,1]*exp(-theta*U[1,1])
    der_t2 = U[1,2]*exp(-theta*U[1,2])
    #score function
    ScF[i,1] = (1/theta+q/(1-q)-(U[1,1]+U[1,2]))-2/((1-q)-t1*t2)*(q-der_t1*t1-der_t2*t2)

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```

IM[i,1] = fdjac(theta,der_LLfr(theta, U))
    vSt[i,1] = ScF*1/sqrt(IM)
    thetag[i,1] = omega + A*vSt[i,1] + B * thetag[i-1,1]
endloop --quiet
return LL
end function

function matrix cfrank_gas_par (matrix *parms, matrix *UU)
    matrix U = {UU[1,1],UU[1,2]}
    matrix LL = zeros(rows(UU),1)
    matrix vSt = zeros(rows(UU),1)
    matrix thetag = zeros(rows(UU),1)
    omega = parms[1,1]
    A = parms[2,1]
    B = parms[3,1]
    thetag[1,1] = omega
    thetat = thetag[1,1]
    omega = omega*(1-B)
    theta = (thetag[1,1])
    expu = exp(theta.*U)
    if abs(theta) < 1e-5
        LL[1,1] = 2*theta*(prodr(U-.5))
    else
        LL[1,1] = log(-theta.*(exp(-theta)-1)) + theta.*sumr(U)
        LL[1,1] = LL[1,1] -2*log(abs(1+exp(theta.*(sumr(U)-1))-sumr(expu)))
    endif
    q = exp(-theta)
    t1 = (1-exp(-theta*U[1,1]))
    t2 = (1-exp(-theta*U[1,2]))
    der_t1 = U[1,1]*exp(-theta*U[1,1])
    der_t2 = U[1,2]*exp(-theta*U[1,2])
    matrix ScF = ones(rows(UU),1)
    ScF = (1/theta+q/(1-q)-(U[1,1]+U[1,2]))-2/((1-q)-t1*t2)*(q-der_t1*t1-der_t2*t2)
    matrix IM = ones(rows(UU),1)
    IM[1,1] = fdjac(theta,der_LLfr(theta, U))
    vSt[1,1] = ScF*1/sqrt(IM)

```

```

    thetag[1,1] = omega + A*vSt[1,1] + B * thetat
loop i = 2..rows(UU) --quiet
    theta = thetag[i-1,1]
    U = {UU[i,1],UU[i,2]}
    expu = exp(theta.*U)
    if abs(theta) < 1e-5
        LL[i,1] = 2*theta*(prodr(U-.5))
    else
        LL[i,1] = log(-theta.*(exp(-theta)-1)) + theta.*sumr(U)
        LL[i,1] = LL[i,1] -2*log(abs(1+exp(theta.*(sumr(U)-1))-sumr(expu)))
    endif
    q = exp(-theta)
    t1 = (1-exp(-theta*U[1,1]))
    t2 = (1-exp(-theta*U[1,2]))
    der_t1 = U[1,1]*exp(-theta*U[1,1])
    der_t2 = U[1,2]*exp(-theta*U[1,2])
    #score function
    ScF[i,1] = (1/theta+q/(1-q)-(U[1,1]+U[1,2]))-2/((1-q)-t1*t2)*(q-der_t1*t1-der_t2*t2)
IM[i,1] = fdjac(theta,der_LLfr(theta, U))
    vSt[i,1] = ScF*1/sqrt(IM)
    thetag[i,1] = omega + A*vSt[i,1] + B * thetag[i-1,1]
endloop --quiet
return sumc(LL)
end function

function matrix est_frank_gas(series x, series y, bool emp[0])
    if emp==1
        matrix U = empirical_cdf(x,y)
    else
        matrix U = {x,y}
    endif
    matrix parms = {1; -3; 5}
    set stopwatch
    catch mle plutoo = cfrank_gas(&parms, &U)
        params parms
    end mle --quiet --lbfgs

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```

    omega = parms[1]
    A = parms[2]
    B = parms[3]
    parmsr = restrpar(omega, A, B)
    err =$error
    matrix s = zeros(1,9)
    gas = cfrank_gas_par(&parmsr,&U)
    s[1] = gas[2]
    s[2] = NA
    s[3] = -gas[1]
    if err > 0
        s[4:6] = IC1(U,s[3],3)
    else
        s[4:6] = {$aic, $bic, $hqc}
    endif
    return s
end function

# Dynamic Gumbel
function matrix GumbLL_der(matrix theta, matrix U)
    derL1 = gumb_der(theta,U)
    derL2 = 1./(gum_pdf(theta,U))
    matrix derLL = derL1.*derL2
    return derLL
end function

function matrix gumb_der(matrix theta, matrix U)
    t1 = (-ln(U[,1])).^theta
    t2 = (-ln(U[,2])).^theta
    t1t2 = t1+t2

    der1 = -(t1t2).^(1./theta).*(-ln(t1t2))./(theta.^2) + (t1.*ln(-ln(U[,1])) + t2.*ln(-ln(U[,2])))
    der2 = (-2.*(ln(t1t2))./(theta.^2) + (-2+2./theta).*(t1.*ln(-ln(U[,1])) + t2.*ln(-ln(U[,2])))
    matrix der12 = gum_pdf(theta,U) .*(der1+der2)
    der3 = gumbel_cdf(U,theta).*((t1t2).^(-2+2./theta)).*((ln(U[,1]).*ln(U[,2])).^(theta-1))./(
    der4 = (t1t2.^(-1./theta)+(theta-1).*(t1t2.^(-1/theta)).*(ln(t1t2))./(theta.^2) - (t1.*ln(-

```

```

matrix der34 = der3.*der4
matrix der = der12+der34
return der
end function

```

```

function matrix cGumb_gas2(scalar A, scalar B, scalar C, matrix U)
    parms = {A;B;C}
    matrix LL = zeros(rows(U),1)
    matrix vSt = zeros(rows(U),1)
    matrix vrho = zeros(rows(U),1)
    matrix thetag = zeros(rows(U),1)
    parms[1,1] = exp(parms[1,1])+1
    omega = parms[1,1]
    A = parms[2,1]
    B = parms[3,1]
    parms = restrpar(omega,A,B)
    omega = parms[1,1]
    A = parms[2,1]
    B = parms[3,1]
    thetag[1,1] = omega
    thetat = thetag[1,1]
    omega = omega*(1-B)
    matrix theta = ones(1,1).*thetag[1,1]
    UU = U[1,]
    v = -log(abs(UU))
    matrix vv = zeros(rows(v),2)
    loop i = 1..rows(v) --quiet
        vv[i,] = sort(v[i,])
    endloop
    vmin = vv[,1]
    vmax = vv[,2]
    logvv = log(vv)
    nlogC = (vmax.*(1+(vmin./vmax).^theta).^ (1./theta))
    lognlogC = log(nlogC)
    logy = log(theta - 1+nlogC) - nlogC + sumr((theta-1).*logvv + v) + (1-2*theta).*lognlogC

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```

LL[1,1] = logy

# score function-----

matrix ScF = ones(rows(U),1)
ScF[1,1] = GumbLL_der(theta, UU)

matrix IM = ones(rows(U),1)
IM[1,1] = fdjac(theta,GumbLL_der(theta,UU))

# 0caled score fun/ction-----
vSt[1,1] = ScF[1,1] * 1/sqrt(abs(IM[1,1]))
# GAS recursion-----
thetag[1,1] = omega + A*vSt[1,1] + B * thetat

UUU = U
loop i = 2..rows(UUU) --quiet
    #theta = exp(thetag[i-1,1])+1
    theta = thetag[i-1,1]
    UU = {UUU[i,1],UUU[i,2]}
    v = -log(abs(UU))
    matrix vv = zeros(rows(v),2)
    loop j = 1..rows(v) --quiet
        vv[j,] = sort(v[j,])
    endloop
    vmin = vv[,1]
    vmax = vv[,2]
    logvv = log(vv)
    nlogC = (vmax.*(1+(vmin./vmax).^theta).^(1./theta))
    lognlogC = log(nlogC)
    logy = log(theta - 1+nlogC) - nlogC + sumr((theta-1).*logvv + v) + (1-2*theta).*lognlogC
    #logy = log(theta - 1+nlogC) - nlogC + sumr((thetam1).*logvv + v) + (1-2*theta).*lognlogC
    LL[i,1] = logy
    ScF[i,1] = GumbLL_der(theta, UU)
    IM[i,1] = fdjac(theta,GumbLL_der(theta,UU))
    vSt[i,1] = ScF[i,1] * 1/sqrt(abs(IM[i,1]))

```

```

        thetag[i,1] = omega + A*vSt[i,1] + B * thetag[i-1,1]
    endloop --quiet
    eval {omega,A,B}
    return LL
end function

function matrix cGumb_gas_par2 (matrix *parms, matrix *U)
    matrix LL = zeros(rows(U),1)
    matrix vSt = zeros(rows(U),1)
    matrix vrho = zeros(rows(U),1)
    matrix thetag = zeros(rows(U),1)

    omega = parms[1,1]
    A = parms[2,1]
    B = parms[3,1]

    thetag[1,1] = omega
    thetat = thetag[1,1]
    omega = omega*(1-B)
    matrix theta = ones(1,1).*thetat[1,1]
    UU = U[1,]
    v = -log(abs(UU))
    matrix vv = zeros(rows(v),2)
    loop i = 1..rows(v) --quiet
        vv[i,] = sort(v[i,])
    endloop
    vmin = vv[,1]
    vmax = vv[,2]
    logvv = log(vv)
    nlogC = (vmax.*(1+(vmin./vmax).^theta).^(1./theta))
    lognlogC = log(nlogC)
    logy = log(theta - 1+nlogC) - nlogC + sumr((theta-1).*logvv + v) + (1-2*theta).*lognlogC
    LL[1,1] = logy

# score function-----

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```

ScF = GumbLL_der(theta, UU)

# derivate seconde della log(pdf)-----
matrix IM = fdjac(theta,GumbLL_der(theta,UU))

#Scaled score fun/ction
vSt[1,1] = ScF[1,1] * 1/sqrt(abs(IM[1,1]))
#GAS recursion
thetag[1,1] = omega + A*vSt[1,1] + B * thetat

UUU = U
loop i = 2..rows(UUU) --quiet
    theta = thetag[i-1,1]
    U = {UUU[i,1],UUU[i,2]}
    v = -log(abs(U))
    matrix vv = zeros(rows(v),2)
    loop j = 1..rows(v) --quiet
        vv[j,] = sort(v[j,])
    endloop
    vmin = vv[,1]
    vmax = vv[,2]
    logvv = log(vv)
    nlogC = (vmax.*(1+(vmin./vmax).^theta).^(1./theta))
    lognlogC = log(nlogC)
    logy = log(theta - 1+nlogC) - nlogC + sumr((theta-1).*logvv + v) + (1-2*theta).*lognlogC
    LL[i,1] = logy

# score function-----

ScF = GumbLL_der(theta, U)
#---derivate seconde della log(pdf)-----

IM = fdjac(theta,GumbLL_der(theta,U))

#Scaled score fun/ction
vSt[i,1] = ScF * 1/sqrt(abs(IM))

```

```

        #GAS recursion
        thetag[i,1] = omega + A*vSt[i,1] + B * thetag[i-1,1]
    endloop --quiet
    return sumc(LL)^theta #~ thetag #-LL[rows(U),1]/rows(U)
end function

function matrix est_gumb_gas2(series x,series y, bool emp[0])
    if emp==1
        matrix U = empirical_cdf(x,y)
    else
        matrix U = {x,y}
    endif
    matrix parms = {1.5; -3; 5}
    parms[1]=log(parms[1]-1)
    catch mle plutoo = cGumb_gas2(&parms, &U)
    params parms
    end mle --quiet --lbfgs
    omega = exp(parms[1])+1
    A = parms[2]
    B = parms[3]
    parmsr = restrpar(omega, A, B)
    err =$error
    matrix s = zeros(1,9)
    gas = cGumb_gas_par2(&parmsr,&U)
    s[1] = gas[2]
    s[2] = NA
    s[3] = -gas[1]
    if err > 0
        s[4:6] = IC1(U,s[3],3)
    else
        s[4:6] = {$aic, $bic, $hqc}
    endif
    return s
end function

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```
# Rotated Dynamic Gumbel

function matrix est_rgumb_gas2(series x,series y, bool emp[0])
  if emp==1
    matrix U = empirical_cdf(x,y)
  else
    matrix U = {x,y}
  endif
  matrix parms = {1.5; -3; 5}
  parms[1]=log(parms[1]-1)
  U = 1-U
  catch mle plutoo = cGumb_gas2(&parms, &U)
  params parms
end mle --quiet --lbfgs
omega = exp(parms[1])+1
A = parms[2]
B = parms[3]
parmsr = restrpar(omega, A, B)
err =$error
gas = cGumb_gas_par2(&parmsr,&U)
matrix s = zeros(1,9)
s[1] = gas[2]
s[2] = NA
s[3] = -gas[1]
#if err > 0
s[4:6] = IC1(U,s[3],3)
return s
end function

#Print Function
function matrix Coprint (matrix Cop, scalar type, matrix mat)
  string start1=obslabel(mat[1,3])
  string start2=obslabel(mat[1,5])
  string end1=obslabel(mat[1,4])
  string end2=obslabel(mat[1,6])
  if start1=start2 && end1=end2
```

```

print " "
    printf "series %d: %s\n", 1, varname(1)
    printf "series %d: %s\n", 2, varname(2)
    print " "
    printf "using observations %s:%s (T = %g)\n",start1,end1,mat[1,1]
else
    print "Time series do not match: "
    printf " 1st series ranges from %s to %s (T=%g)\n",start1,end1,mat[1,1]
    printf " 2nd series ranges from %s to %s (T=%g)\n",start2,end2,mat[1,2]
endif
print " "
if type = 2
    matrix coef = {Cop[1],Cop[9];Cop[2],Cop[10]}
    matrix stat = {Cop[3];Cop[4];Cop[5];Cop[6];Cop[7];Cop[8]}
    string name = "Parameter,dof,Log-Lik, AIC, BIC, HQC, Upper Tail-Dep, Lower Tail-Dep"
    modprint coef name stat
else
    matrix coef = {Cop[1],Cop[9]}
    matrix stat = {Cop[3];Cop[4];Cop[5];Cop[6];Cop[7];Cop[8]}
    string name = "Parameter,Log-Lik, AIC, BIC, HQC, Upper Tail-Dep, Lower Tail-Dep"
    modprint coef name stat
endif
return coef
end function

## Main Function-----

function matrix BiCopulaEst(series x, series y, scalar type, scalar emp, bool pr)
matrix mat=match(x,y)
if type=1
    if mat[1,1] > 0
        Cop=NormalCopula_es(x,y,emp)
        if pr = 1
            print " "
            print "-----"

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```
        print "BIVARIATE NORMAL COPULA ESTIMATION VIA MAXIMUM LIKELIHOOD"
        pi = Coprint(Cop, type,mat)
    endif
else
    return 0
endif

elif type=2
    if mat[1,1] > 0
        Cop=tCopula_es(x,y,emp)
        if pr = 1
            print " "
            print "-----"
            print "BIVARIATE STUDENT-T COPULA ESTIMATION VIA MAXIMUM LIKELIHOOD"
            pi = Coprint(Cop, type,mat)
        endif
    else
        return 0
    endif

elif type=3
    if mat[1,1] > 0
        Cop = ClaytonCopula_es(x,y,emp)
        if pr = 1
            print " "
            print "-----"
            print "BIVARIATE CLAYTON COPULA ESTIMATION VIA MAXIMUM LIKELIHOOD"
            pi = Coprint(Cop, type,mat)
        endif
    else
        return 0
    endif

elif type==4
    if mat[1,1] > 0
```

```

Cop = RClaytonCopula_es(x,y,emp)
if pr = 1
    print " "
    print "-----"
    print "BIVARIATE ROTATED-CLAYTON COPULA ESTIMATION VIA MAXIMUM LIKELIHOOD"
    pi = Coprint(Cop, type,mat)
endif
else
    return 0
endif

elif type==5
    if mat[1,1] > 0
        Cop = FrankCopula_es(x,y,emp)
        if pr = 1
            print " "
            print "-----"
            print "BIVARIATE FRANK COPULA ESTIMATION VIA MAXIMUM LIKELIHOOD"
            pi = Coprint(Cop, type,mat)
        endif
    else
        return 0
    endif

elif type==6
    if mat[1,1] > 0
        Cop = GumbelCopula_es(x,y,emp)
        if pr = 1
            print " "
            print "-----"
            print "BIVARIATE GUMBEL COPULA ESTIMATION VIA MAXIMUM LIKELIHOOD"
            pi = Coprint(Cop, type,mat)
        endif
    else
        return 0
    endif

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```
elif type==7
  if mat[1,1] > 0
    Cop = RGumbelCopula_es(x,y,emp)
    if pr = 1
      print " "
      print "-----"
      print "BIVARIATE ROTATED-GUMBEL COPULA ESTIMATION VIA MAXIMUM LIKELIHOOD"
      pi = Coprint(Cop, type,mat)
    endif
  else
    return 0
  endif
elif type ==8
  if mat[1,1] > 0
    Cop = est_norm_gas(x,y,emp)
    if pr = 1
      print " "
      print "-----"
      pi = Coprint(Cop,type,mat)
    endif
  else
    return 0
  endif

elif type==9
  if mat[1,1] > 0
    Cop = est_clt_gas(x,y,emp)
    if pr = 1
      print " "
      print "-----"
      pi = Coprint(Cop,type,mat)
    endif
  else
    return 0
  endif
```

```

elif type ==10
    if mat[1,1] > 0
        Cop = est_rclt_gas(x,y,emp)
        if pr = 1
            print " "
            print "-----"
            pi = Coprint(Cop,type,mat)
        endif
    else
        return 0
    endif

elif type ==11
    if mat[1,1] > 0
        Cop = est_frunk_gas(x, y, emp)
        if pr = 1
            print " "
            print "-----"
            pi = Coprint(Cop,type,mat)
        endif
    else
        return 0
    endif

elif type ==12
    if mat[1,1] > 0
        Cop = est_gumb_gas(x,y,emp)
        if pr = 1
            print " "
            print "-----"
            pi = Coprint(Cop,type,mat)
        endif
    else
        return 0
    endif

```

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

```
elif type ==13
    if mat[1,1] > 0
        Cop = est_rgumb_gas(x,y,emp)
        if pr = 1
            print " "
            print "-----"
            pi = Coprint(Cop,type,mat)
        endif
    else
        return 0
    endif
elif type ==14
    if mat[1,1] > 0
        Cop = est_t_gas(x,y,emp)
        if pr = 1
            print " "
            print "-----"
            pi = Coprint(Cop,type,mat)
        endif
    else
        return 0
    endif
endif
return Cop

end function
```

Lista titoli utilizzati

<u>Azienda</u>	<u>Simbolo</u>	<u>num. nodo</u>
ATLANTIA	ATL	1
AZIMUT HOLDING	AZM	2
BPER BANCA	BPE	3
BUZZI UNICEM	BZU	4
CAMPARI	CPR	5
ENI	ENI	6
FIAT CHRYSLER AUTOM.	FCA	7
GENERALI	G	8
INTESA SAN PAOLO	ISP	9
MEDIASET	MS	10
MEDIO BANCA	MB	11
PRYSMIAN	PRY	12
RECORDATI	REC	13
SAIPEM	SPM	14
STMICROELECTRONICS	STM	15
TENARIS	TEN	16
TELECOM ITALIA	TIT	17

Statistiche descrittive

Tabella 5.0.1: Statistiche descrittive, utilizzando le osservazioni 2010-01-01–2015-10-01

Variabile	Media	Mediana	Minimo	Massimo
V1	-0,00043906	-0,00045840	-0,080786	0,095349
V2	-0,00057777	-0,00053695	-0,096849	0,14793
V3	0,00026450	0,00046139	-0,15005	0,28251
V4	-0,00061004	-0,00058944	-0,076879	0,099483
V5	-0,00058523	-0,00088075	-0,076588	0,078738
V6	-0,00016739	0,00000	-0,067425	0,096402
V7	-0,00050576	-0,0012154	-0,15186	0,17604
V8	-8,5702e-005	0,00000	-0,089403	0,18354
V9	-0,00032044	-0,00072418	-0,13493	0,26060
V10	-0,00017215	-0,00056743	-0,094745	0,23847
V11	-4,0014e-005	0,00000	-0,27658	0,18835
V12	-0,00039854	-0,00082529	-0,085374	0,13366
V13	-0,0011234	-0,0012908	-0,063106	0,092941
V14	0,0011710	0,00000	-0,26480	0,41985
V15	-0,00064191	-0,00033556	-0,11015	0,15383
V16	-0,00011597	0,00000	-0,13653	0,14495
V17	2,2777e-005	0,00055252	-0,089435	0,18172

Variabile	Dev. Std.	Coeff. di variazione	Asimmetria	Curtosi
V1	0,017345	39,504	0,10847	2,1831
V2	0,024140	41,781	0,23475	2,1983
V3	0,034341	129,83	0,16429	4,2048
V4	0,023302	38,198	0,16549	0,95613
V5	0,014963	25,567	0,26626	2,2342
V6	0,016731	99,953	0,16714	2,4145
V7	0,028176	55,711	0,32073	2,5925
V8	0,019843	231,53	0,43327	6,7121
V9	0,028938	90,306	0,70007	6,7340
V10	0,026324	152,91	0,63007	5,6056
V11	0,029741	743,28	-0,54917	8,5194
V12	0,021107	52,962	0,28564	2,1142
V13	0,015782	14,048	0,21555	2,1296
V14	0,033182	28,336	2,3540	34,116
V15	0,025117	39,129	0,30056	3,3897
V16	0,020995	181,04	-0,0068913	4,1744
V17	0,023965	1052,1	0,22313	3,0038

Bibliografia

- AAS, K., C. CZADO, A. FRIGESSI, AND H. BAKKEN (2007): “Pair-copula constructions of multiple dependence,” *Insurance: Mathematics and Economics*.
- ALEXANDER, C. (2008): *Market Risk Analysis*, vol. 3. Wiley & Sons , London.
- ALEXANDER, C., AND A. CHIBUMBA (1996): “Multivariate orthogonal factor GARCH,” .
- ALEXANDER, G. J., AND A. M. BAPTISTA (2008): “Active portfolio management with benchmarking: adding a value-at-risk constraint,” *Journal of Economic Dynamics & Control*, 32, 779–820.
- (2010): “Active portfolio management with benchmarking: a frontier based on alpha,” *Journal of Banking & Finance*, 34, 2185–2197.
- ALMEIDA, C., AND C. CZADO (2012): “Bayesian inference for stochastic time-varying copula models,” *Computational Statistics & Data Analysis*, 56, 1511–1527.
- ALMEIDA, C., C. CZADO, AND H. MANNER (2016): “Modeling high-dimensional time-varying dependence using dynamic D-vine models,” *Applied Stochastic Models in Business and Industry*, 32, 621–638.
- ANG, A., AND G. BEKAERT (2002): “International asset allocation with regime shifts,” *Review of Financial Studies*, 15, 1137–1187.
- ARDITTI, T. (1967): “Risk and the Required Return on Equity,” *Journal of Finance*, 22, 19–36.
- BAILLIE, R. T., T. BOLLERSLEV, AND H. O. MIKKELSEN (1996): “Fractionally integrated generalized autoregressive conditional heteroskedasticity.,” *Journal of Econometrics*, 74, 3–30.

ASSET ALLOCATION E COPULAE MULTIVARIATE DINAMICHE

- BAUWENS, L., AND S. LAURENT (2005): “A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models,” *Journal of Business & Economic Statistics*, 23, 346–354.
- BEDFORD, T., AND R. COOKE (2001): “Probability density decomposition for conditionally dependent random variables modeled by vines,” *Annals of Mathematics and Artificial Intelligence*, 32, 245–268.
- (2002): “Vines - a new graphical model for dependent random variables,” *Annals of Statistics*, 30, pp. 1031–1068.
- BEKAERT, G., AND C. HARVEY (1995): “Time-varying world market integration,” *Journal of Finance* 50, 2, 403–444.
- BERG, D., AND K. AAS (2007): “Model for construction of multivariate dependence, Technical report,” .
- BERG, D., AND K. ASS (2009): “Models for construction of higher-dimensional dependence: A comparison study,” *Journal of Finance*, 15, 639–659.
- BICKEL, P., AND K. DOKSUM (2007): *Mathematical Statistics: Basic Ideas and selected Topics*. Pearson Prentice Hall, Uppes Saddle River.
- BLACK, F., AND M. SCHOLES (1973): “The pricing of options and corporate liabilities,” *Journal of Political Economy*, 81, 637–654.
- BOLLERSLEV, T. (1986): “Generalized autoregressive conditional heteroskedasticity,” *Journal of Econometrics*, 31, pp. 307–327.
- (1990): “Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model,” *The Review of Economics and Statistics*, pp. 498–505.
- BOLLERSLEV, T., R. ENGLE, AND J. D. B. F. NELSON (1994): “ARCH Models,” 4.
- BOLLERSLEV, T., R. ENGLE, AND J. M. WOOLDRIDGE (1988): “A capital asset pricing model with time varying covariances,” *Journal of Political Economy*, 96, 116–131.
- BRECHMANN, E. C., AND C. CZADO (2013): “Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50,” *Statistics & Risk Modeling*, 30, 307.342.

- BRECHMANN, E. C., C. CZADO, AND K. AAS (2012): “Truncated regular vines in high dimensions with application to financial data,” *Canadian Journal of Statistics*, 40, 68–85.
- CHAN, L., J. KARCESKI, AND J. LAKONISHOK (1999): “On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model,” *The Review of Financial Studies*, 12, 937–974.
- CHEN, X., AND Y. FAN (2006): “Estimation of copula-based-semiparametric time series models,” *Journal of Econometrics*, 130, 307–335.
- CHERUBINI, U., AND E. LUCIANO (2001): “Value at Risk trade-off and capital allocation with copulas,” *Economic Notes*, 30, 235–256.
- CHERUBINI, U., E. LUCIANO, AND W. VECCHIATO (2004): *Copula Methods in Finance*. Chichester: John Wiley & Sons.
- CHOLLETE, L., A. HEINEN, AND A. VALDESOGO (2009): “Modeling international financial returns with a multivariate regime-switching copula,” *Journal of Financial Economy*, 7, 437–480.
- CORMEN, T., C. LEISERSON, R. RIVEST, AND C. STEIN (2001): “Introduction to Algorithms,” in *2nd Edition*. The MIT Press, Cambridge.
- CREAL, D., S. KOOPMAN, AND A. LUCAS (2013): “Generalized autoregressive score models with applications,” *Journal of Applied Econometrics*, 28, 777–795.
- DE LIRA SALVATIERRA, I., AND A. J. PATTON (2015): “Dynamic copula models and high frequency data,” *Journal of Empirical Finance*, 30, 120–135.
- DIEBOLD, F. X., J. HAHN, AND A. S. TAY (1999): “Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High Frequency Return on Foreign Exchange,” *Review of Economics and Statistics*, 81, 661–673.
- DISSMANN, J., E. C. BRECHMANN, C. CZADO, AND D. KUROWICKA (2013): “Selecting and estimating regular vine copulae and application to financial returns,” *Computational Statistics & Data Analysis*, 59, 52–69.
- EFRON, B., AND D. HINKLEY (1978): “Assessing the accuracy of the maximum likelihood estimator: Observed versus expected fisher information,” *Biometrika*, 65 (3), 457–482.

- EIKE, C., BRECHMANN, AND C. CZADO (2013): “Risk Management with High-Dimensional Vine Copulas: An Analysis of the Euro Stoxx 50,” 30.
- EMBRECHTS, P., H., AND A. JURI (2003): “Using Copulae to bound the Value-at-Risk for functions of dependent risk,” *Finance & Stochastics*, 7, 145–167.
- ENGLE, R., AND K. SHEPPARD (2001): “Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH,” *NBER Working Paper 8554*.
- ENGLE, R. F. (1982): “Autoregressive conditional heteroskedasticity with estimates of the U.K inflation,” *Econometrica*, 50, pp. 987–1008.
- (2002): “Dynamic conditional correlation: a simple class of multivariate GARCH models,” *Journal of Business and Economic Statistics*, 20, 339–350.
- ENGLE, R. F., AND K. F. KRONER (1995): “Multivariate simultaneous generalized ARCH,” *Econometric Theory*, 11, 122–150.
- ENGLE, R. F., D. LILIEN, AND R. ROBINS (1987): “Estimating time varying risk premia in the term structure: The ARCH-M model,” *Econometrica*, 55, 391–407.
- ENGLE, R. F., NG., AND M. ROTHSCHILD (1990): “Asset pricing with a factor ARCH covariance structure: Empirical estimates fo Treasury Bills,” *Journall of Econometrics*, 45, 213–238.
- ENGLE, R. F., AND J. RUSSELL (1998): “Autoregressive conditional duration: a new model for irregularly spaced transaction data.,” *Econometrica*, 66, 1127–1162.
- FANG, Y., L. MADSEN, AND L. LIU (2014): “Comparison of Two Methods to Check Copula Fitting,” 44, 53–61.
- FISHER, R. A. (1932): “Statistical Methods for Research Workers,” .
- GENEST, C., AND R. B. (2008): “Validity of the Parametric Bootstrap for Goodness-of-fit Testing in Semiparametric Models,” *Annales de Institut Henri Poincaré*, 44, 1096–1127.
- GENEST, C., K. GHOUDI, AND L. RIVEST (1993): “A semiparametric estimation procedure of dependence parameters in multivariate families of distribution,” *Biometrika*, 82, 543–552.
- GENEST, C. R. B., AND B. D. (2009): “Goodness-of-fit Test for Copulas: A Review and a Power Study,” *Insurance: Mathematics and Economics*, 44, 199–214.

- GLOSTEN, L. R., R. JAGANNATHAN, AND D. E. RUNKLE (1993): “On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks,” *Journal of Finance*, 48, 1779–1801.
- GODAMBE, V. (1960): “An optimim property of regular maximum likelihood estimation,” *Ann. Math. Statist.*, 31, 1208–1211.
- GONZALO-RIVERA, G. (1998): “Smooth transition GARCH models,” *Studies in Nonlinear Dynamics and Econometrics.*, 3, 61–78.
- GUIDOLIN, M., AND A. TIMMERMANN (2008): “International asset allocation under regime switching, skew, and kurtosis preferences,” *Review of Financial Studies*, 21, 889–935.
- HAFF, I. H. (2013): “Estimating the parameters of a pair copula construction,” *Bernoulli*, 19(2), 462–491.
- HAFNER, C. M., AND H. MANNER (2012): “Dynamic stochastic copula models: Estimation, inference and applications,” *Journal of Applied Econometrics*, 7, 269–295.
- HAMAQ, Y., R. MASULIS, AND V. NG (1990): “Correlations in price changens and volatility across international stock markets,” *Review of Financial Studies* 3 (2), pp. 281–307.
- HAMILTON, J. (1989): “A new approach to the economic analysis of nonstationary time series and the business cycle,” *Econometrica*, 57, 357–384.
- HANSEN, B. (1994): “Autoregressive condiitonal density estimation,” *International Economic Review*, 35, 705–730.
- HARVEY, C., J. LIECHTY, M. LIECHTY, AND P. MULLER (2010): “Portfolio selection with higher moments,” *Quantitative Finance*, 10, 469–485.
- HEINEN, A., AND A. VALDESOGO (2008): “Asymmetric CAPM dependence for large dimensions: the Canonical Vine Autoregressive Model,” Working paper.
- HIGGINS, M. L., AND A. K. BERA (1992): “Class of Nonlinear ARCH Models.,” *International Economic Review*, 33, 137–158.
- HOBÆK HAFF, I. (2013): “Parameter estimation for pair-copula constructions.,” *Bernoulli*, 19 (2), 462–491.
- JOE, H. (1996): “Families of m-variate distributions with given margins and $(m-1)/2$ bivariate dependence parameters,” .

- (1997): “Multivariate Models and Dependence Concepts,” .
- JOE, H., AND J. J. XU (1996): “The estimation method of inference functions for margins for multivariate models,” Working paper.
- JONDEAU, E., AND M. ROCKINGER (2003): “Conditional volatility, skewness, and kurtosis: existence, persistence and comovements,” *Journal of Economics Dynamics and Control*, 27, 1699–1737.
- (2005): “Conditional Asset Allocation under Non-Normality: How Costly is the Mean-Variance Criterion?,” *HEC Lausanne and FAME, Institute of Banking and Finance, Switzerland*, Working paper.
- (2006b): “Optimal Portfolio Allocation under Higher Moments,” *European Financial Management*, 12(1), 29–55.
- (2012): “On the importance of time variability in higher moments for asset allocation.,” *Journal of Financial Economy*, 10, 84–123.
- JORION, P. (2003): “Portfolio optimization with constraints on tracking error,” *Financial analysts journal*, 59, 70–82.
- KINOSHITA, R. (2015): “Asset allocation under higher moments with the GARCH filter,” *Empirical Economics*, 49, 235–254.
- KOJADINOVIC, I., AND J. YAN (2011): “A Goodness-of-fit Test for Multivariate Multiparameter Copulas Based on Multiplier Central Limit Theorems,” *Statistics and Computing*, 21, 17–30.
- KRONER, K., AND V. NG (1998): “Modeling asymmetric comovements of asset returns,” *Review of Financial Studies* 11, 4, 817–844.
- KUROWICKA, D., AND R. COOKE (2006): “Uncertainty Analysis with High Dimensional Dependence Modelling,” .
- LINDSAY, B. (1988): “Composit likelihood methods,” *Contemporary Mathematics*, 80, 220–239.
- LONGIN, F., AND B. SOLNIK (2001): “Extreme correlation of international equity markets,” *Journal of Finance*, 56, 649–676.

- MANNER, H., AND J. SERGERS (2011): “Tails of correlations mixtures of elliptical copulas,” *Insurance Math. Econom.*, 48, 153–160.
- MARKOWITZ, H. (1952): “Portfolio selection,” *Journal of Finance*, 7, 77–91.
- MCNEIL, A.J. FREY, R., AND P. EMBRECHTS (2005): “Quantitative Risk Management: Concepts, Techniques and Tools,” .
- MORALES-NÁPOLES, O. (2008): “Bayesian belief nets and vines in aviation safety and other applications,” *Ph.D. thesis*.
- NELSON, D. B. (1990): “Stationarity and persistence in the GARCH(1,1) model,” *Econometric Theory*, 6, 318–334.
- OKIMOTO, T. (2008): “New evidence of asymmetric dependence structure in international equity markets,” *J. Finan. Quant. Anal*, 43, 787–815.
- PALOMBA, G., AND L. RICCETTI (2012): “Portfolio frontiers with restrictions to tracking error volatility and value at risk,” *Journal of Banking & Finance*, 36, 2604–2615.
- PATTON, A. (2004): “On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation,” *Journal of Financial Econometrics*, 2(1), 130–168.
- (2006): “Estimation of multivariate copula models for time series of possibly different lengths,” *Journal of Applied Econometrics*, 21 (2), 147–173.
- QUENOUILLE, M. H. (1949): “Problems in Plane Sampling,” *The Annals of Mathematical Statistics*, 20, 355–375.
- RICCETTI, L. (2010): “The Use of Copulas in Asset Allocation. When and How a Copula Model can be Useful?,” .
- (2013): “A copula–GARCH model for macro asset allocation of a portfolio with commodities,” *Empirical Economics*, 44, 1315–1336.
- RICHARD, J. F., AND W. ZHANG (2007): “Efficient high-dimensional importance sampling,” *Journal of Econometrics*, 141, 1385–1411].
- RÉMILLARD, B. (2010): “Goodness-of-fit tests for copulas of multivariate time series,” *Working Paper*.
- RODRIGUEZ, J. (2007): “Measuring financial contagion: a copula approach,” *Journal of Empirical Finance*, 14, 40–423.

- ROLL, R. (1992): “A mean/variance analysis of tracking error,” *Journal of portfolio management*, 18, 13–22.
- ROSENBERG, J. V., AND T. SCHUERMAN (2006): “A general approach to integrated risk management with skewed, fat-tailed risks,” *Journal of Financial Economics*, 79, 569–514.
- SAHU, S., D. DEY, AND M. BRANCO (2003): “A new class of multivariate skew distributions with applications to bayesian regression models,” *The Canadian Journal of Statistics*, 31, 129,150.
- SAMUELSON, P. (1970): “The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Higher Moments,” *Review of Economic Studies*, 37, 537–542.
- SCHEFFER, M., AND G. WIESS (2015): “Mixture pair-copula-constructions,” 54, 175–191.
- SCHEPSMEIER, U., AND J. STÖBER (2014): “Derivatives and Fisher information of bivariate copulas,” 55.
- SERLING, M. (1980): “Approximation Theorems of Mathematical Statistics,” .
- SHIN, J. H., AND T. A. LOUIS (1995): “Inference on the association parameter in copula models for bivariate survival data,” *Biometrics*, 51, 229–231.
- SKLAR, A. (1959): “Fonctions de répartition à n dimensions et leurs marges,” 8, 229–231.
- SUSMEL, R., AND R. ENGLE (1994): “Hourly volatility spillovers between international equity markets,” *Journal of International Money and Finance* 13, 1, 3–25.
- TSE, Y., AND A. TSUI (2002): “A multivariate generalized autoregressive conditional heteroschedasticity model with time-varying correlations,” *Journal of Business and Economic Statistics* 20, 3, 351–362.
- TSUKAHARA, H. (2005): “Semiparametric estimation in copula models,” *Canad. J. Statist.*, 33, 357–375.
- VAN DER VAART, A., AND J. WELLNER (1996): *Multiplier Central Limit Theorems*, vol. Weak Convergence and Empirical Processes. Springer.
- VAN DER WEIDE, R. (2002): “GO-GARCH: a multivariate generalized orthogonal GARCH model,” *Journal of Applied Econometrics*, 17, 549–564.

- VECCHIA, A. V. (1988): “Estimation and model identification for continuous spatial processes,” *Journal of the Royal Statistical Society. Series B*, 50 (2), 297–312.
- VON NEUMANN, J., AND O. MORGENSTERN (1947): *Theory of games and economic behaviour*, vol. Princeton, New Jersey. Princeton University Press.
- VUONG, Q. H. (1989): “Likelihood ratio test for model selection and non-nested hypotheses,” *Econometrica*, 57, 307–333.
- WHITE, H. (1994): “Estimation, Inference and Specification Analysis,” *Cambridge University Press*.
- ZAKOIAN, J. M. (1994): “Threshold heteroskedastic models,” *Journal of Economic Dynamic and Control*, 18, pp. 931–955.