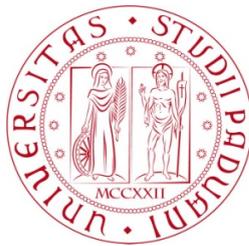


Università degli studi di Padova
Dipartimento di Scienze Statistiche
Corso di Magistrale in
Scienze statistiche



TESI DI LAUREA

**INFERENZA BAYESIANA PER MODELLI AUTOREGRESSIVI
VETTORIALI QUANTILICI CON APPLICAZIONE AL CALCOLO
DEL RISCHIO SISTEMICO**

Relatore Prof. Mauro Bernardi
Dipartimento di Scienze Statistiche

Laureando Piergiacomo Andrea Carlesi
Matricola N 1238856

Anno Accademico 2021/2022

Indice

Introduzione	7
1 Regressione quantilica	9
1.1 Il problema della regressione quantilica	9
1.2 Regressione quantilica bayesiana	12
1.2.1 Introduzione modellazione bayesiana	12
1.2.2 Modellazione quantilica bayesiana	13
1.2.3 Estensione al caso multidimensionale	16
2 Il modello autoregressivo vettoriale	19
2.1 Il modello VAR(p)	19
2.1.1 Stazionarietà	21
2.2 Il modello autoregressivo vettoriale con restrizioni	22
2.3 Analisi Strutturali	23
2.3.1 <i>Impulse response function</i>	23
2.3.2 <i>Forecast error variance decomposition</i>	26
2.4 VAR bayesiano	27
2.5 <i>Spike-and-slab</i>	28
3 Modello VAR quantilico	31
3.1 Modelli VAR quantilici in ambito frequentista	31
3.2 Modello VAR quantilico bayesiano	33
3.2.1 Generazione unica della matrice di covarianza	39
3.3 VAR quantilico con restrizioni	40
3.4 Estensione <i>spike-and-slab</i>	43
3.5 Analisi Strutturali quantiliche	48

3.5.1	Impulse response function quantiliche	49
3.5.2	FEVD quantilica	51
4	Algoritmi di stima	55
4.1	Introduzione agli algoritmi MCMC	55
4.2	Algoritmi di stima	57
4.2.1	Algoritmo di stima modello VAR quantilico	58
4.2.2	Algoritmo per generazione unica della matrice di covarianza	59
4.2.3	Algoritmo di stima VAR con restrizioni	60
4.2.4	Algoritmo di stima <i>spike-and-slab</i>	61
4.3	Esempi con dati simulati	64
4.3.1	Strumenti di valutazione	64
4.3.2	Studio dei dati simulati	66
4.3.3	Osservazioni finali	72
5	Applicazione su dati reali	77
5.1	Misure di rischio	77
5.2	Dataset <i>CDS</i>	79
5.3	Analisi Italia, Grecia, Germania	81
5.3.1	Analisi strutturali quantiliche	85
5.4	Analisi <i>CDS</i> post-crisi	89
5.4.1	Modelli stimati	89
5.4.2	Analisi strutturali	93
5.5	Conclusione dell'analisi	95
	Conclusioni	105
	Bibliografia	107
	Appendice A	111
	Appendice B	117

Elenco delle tabelle

4.1	Valutazione tramite test LRcc delle capacità di stima dei quantili per le simulazioni con $d = 4$	67
4.2	Tabelle riassuntiva delle serie generate per \mathbf{D} e β , con dimensione $d = 4$	68
4.3	Tabelle riassuntiva delle serie generate per Σ e per i parametri relativi al modello con priori <i>spike-and-slab</i> , dimensione $d = 4$	68
4.4	Confronto tra probabilità a posteriori di essere generato dalla <i>slab</i> e vero valore del parametro con cui si è generata la distribuzione.	69
4.5	Valutazione tramite test LRcc delle capacità di stima dei quantili per le simulazioni con $d = 6$	70
4.6	Tabelle riassuntiva delle serie generate per \mathbf{D} e β , dimensione $d = 6$	71
4.7	Tabelle riassuntiva delle serie generate per Σ e per i parametri della <i>spike-and-slab</i> , dimensione $d = 6$	71
5.1	<i>Potential scale reduction factors</i> multivariato per le catene MCMC generate.	82
5.2	Matrici di covarianza stimate, generate tramite algoritmo MCMC.	83
5.3	<i>generalised quantile forecast error variance decomposition</i> , 4 settimane, quantile 0.1.	87
5.4	<i>Generalised quantile forecast error variance decomposition</i> , 4 settimane, quantile 0.5.	87
5.5	<i>Generalised quantile forecast error variance decomposition</i> , 4 settimane, quantile 0.9.	87

5.6	Tabella riassuntiva per la valutazione delle capacità di stima dei modelli proposti tramite test LRcc.	90
5.7	Stime matrice di covarianza, da Σ	92
8	<i>Potential scale reduction factor</i> per β nei quantili 0.1, 0.5, 0.9.	112
9	Effective sample size per β , quantili 0.1, 0.5, 0.9.	113
10	<i>Potential scale reduction factors</i> multivariato per le catene generate nei modelli selezionati.	113
11	<i>Effective sample size</i> media per le catene generate nei modelli selezionati.	116

Introduzione

L'analisi del rischio è un importante tema in ambito economico e finanziario. Una attenzione crescente negli ultimi decenni, anche a seguito delle molteplici crisi economiche che si sono verificate, è stata posta sulla corretta valutazione e stima dei fattori di rischio. Le situazioni di crisi sono spesso frutto di una serie di eventi al di fuori del normale comportamento dei sistemi, risultando ancora più difficili da prevedere. In ambito statistico queste situazioni possono essere interpretate come realizzazioni estreme per le variabili modellate, lontane dai valori medi solitamente osservati. I classici metodi di modellazione statistica, basati sull'analisi del comportamento dei fenomeni in media, non riescono quindi a cogliere le caratteristiche di questi fenomeni ed eventi più rari. La modellazione quantilica è uno strumento in grado di ampliare la conoscenza dei fenomeni di interesse, permettendo una maggiore conoscenza della distribuzione delle variabili e le relazioni tra esse in tutti i quantili della distribuzione. L'obiettivo di questa tesi è quello di ampliare gli strumenti a disposizione per la modellazione delle serie storiche multivariate, utilizzando modelli autoregressivi vettoriali (VAR) quantilici. In letteratura i modelli VAR quantilici sono stati poco approfonditi e lo si è fatto quasi sempre con un approccio frequentista. In questa tesi si è cercato di utilizzare la maggiore flessibilità e le possibilità offerte dall'approccio bayesiano per proporre un modello VAR quantilico, stimabile tramite algoritmi MCMC. Viene proposta una estensione del modello VAR quantilico, basata sull'introduzione di vincoli lineari sulla struttura dei coefficienti, struttura che permette una maggiore elasticità nella stima e nella definizione di modelli di tipo VAR. Viene successivamente introdotto l'utilizzo di priori *spike-and-slab* nel modello VAR quantilico, per introdurre un meccanismo di valutazione dell'importanza delle

covariate e dei ritardi autoregressivi. Vengono successivamente proposte delle versioni delle analisi strutturali, tecniche specifiche per l'analisi dei modelli VAR in grado di aumentare l'interpretabilità delle serie modellate, definite specificamente per i modelli VAR quantilici presentati. I metodi e modelli proposti vengono applicati per il calcolo del rischio sistemico, corrispondente a variazioni delle serie sui quantili estremi, modellando le serie dei prezzi dei *Credit Default Swap (CDS)* sui debiti sovrani europei. Nel capitolo 1 viene presentata la modellazione quantilica, mostrando i risultati probabilistici su cui si fondano i modelli presentati. Nel capitolo 2 vengono presentati i modelli VAR in media, con relative proprietà e con le analisi strutturali classiche; vengono poi presentate le priori usate per la stima di VAR bayesiani, e le priori *spike-and-slab*. Il capitolo 3 presenta la derivazione del modello VAR quantilico da me proposto, con le relative distribuzioni a posteriori, e la derivazione delle estensioni proposte per il modello stesso. A fine capitolo vengono presentati gli strumenti di analisi strutturale proposti per la modellazione VAR quantilica. Nel capitolo 4 vengono presentati gli algoritmi usati per la stima e viene presentato un breve studio di simulazione con i modelli e algoritmi introdotti. Il capitolo 5 contiene una analisi del rischio sistemico associato ai prezzi dei *Credit Default Swap* sui debiti sovrano europei.

Capitolo 1

Regressione quantilica

In questo capitolo viene brevemente introdotta la storia e lo sviluppo della modellazione quantilica, partendo dall'articolo seminale di Bassett e Koenker 1978 che per primi introdussero il concetto di modellazione quantilica, presentando poi alcuni sviluppi più recenti, in ambito bayesiano, quali Yu e Moyeed 2001 e Kozumi e Kobayashi 2011. Viene infine brevemente mostrato come estendere al caso multidimensionale la modellazione proposta da Kozumi e Kobayashi 2011, così da ottenere la base teorica su cui verranno sviluppati i modelli introdotti nel capitolo 3.

1.1 Il problema della regressione quantilica

Immaginiamo di voler modellare il comportamento di una variabile aleatoria Y condizionatamente ad un insieme di variabili X . Supponiamo inoltre di avere un insieme di osservazioni, realizzazioni di Y e X , y vettore relativo alle osservazioni campionarie di Y , $y \in \mathbb{R}^{n \times 1}$, e definiamo l'insieme di osservazioni relative alle covariate da studiare come \mathbf{X} , matrice $\mathbf{X} \in \mathbb{R}^{n \times m}$. Definiamo quindi come y_i l' i -esimo elemento della variabile risposta da modellare in funzione di \mathbf{x}_i , i -esima riga di \mathbf{X} , vettore contenente i valori osservati per le covariate relative alla i -esima osservazione. Un modo semplice, ed intuitivo, per modellare la relazione che intercorre tra Y e X , alla luce dei dati osservati, è quello di ricorrere ad un modello lineare sulla media di Y , ossia stimare dei coefficienti β tali che $E(Y|X) = X^\top \beta$. La funzione di perdita da minimizzare

per ottenere una regressione sulla media, in grado di fornire una stima di β , è la soluzione ai minimi quadrati $\hat{\beta}$ definita come:

$$\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^n |y_i - \mathbf{x}_i^T \beta|^2. \quad (1.1)$$

La modellazione in media, ottenuta come in (1.1), risulta poco robusta, facilmente influenzabile da *outlier* o valori errati. La modellazione sulla mediana di Y , al contrario, pur dando sempre una modellizzazione di Y "centrata", $P(Y > X^T \beta) = P(Y < X^T \beta) = 0.5$, garantisce una robustezza maggiore; si ottiene una stima di β in mediana, $\hat{\beta}$, come:

$$\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^n |y_i - \mathbf{x}_i^T \beta|. \quad (1.2)$$

Le modellazioni in media, o in mediana, forniscono informazioni solamente sulla relazione che intercorre in un punto centrale della distribuzione di Y , tralasciando tutto il resto della distribuzione, e quindi fornendoci una conoscenza limitata del rapporto che intercorre tra Y e X . Un possibile metodo per accrescere la conoscenza della densità di Y , ed in che modo X ne influenza l'intera distribuzione, è quello di stimare la relazione tra Y e X non solo sulla mediana, ma sui diversi quantili della distribuzione di Y . Ricordo che il generico quantile τ in $[0, 1]$ di una distribuzione Y è un elemento Y_τ appartenente al supporto di Y tale che

$$P(Y \leq Y_\tau) = F_Y(Y_\tau) = \tau,$$

da cui possiamo definire la funzione quantile come;

$$Q_Y(\tau) = F_Y^{-1}(\tau) = \inf(Y : F(Y) > \tau).$$

Bassett e Koenker 1978, partendo dallo stimatore in mediana definito in (1.2) estesero il concetto alla stima e alla modellazione modellazione condizionata nei quantili di Y . Bassett e Koenker 1978 proposero la seguente funzione di perdita, per effettuare una regressione sul quantile τ , ossia modellando la distribuzione condizionata della variabile risposta Y nel suo quantile τ invece che in mediana. La funzione di perdita, viene così ottenuta pesando gli

scarti assoluti del modello per $1 - \tau$ se il valore predetto è inferiore al valore osservato, e per τ nel caso in cui sia maggiore. Se prendiamo ad esempio $\tau = 0.7$, si avrà che gli scarti, $|y_i - \mathbf{x}_i^\top \boldsymbol{\beta}| > 0$ verranno moltiplicati per 0.7, mentre gli scarti $|y_i - \mathbf{x}_i^\top \boldsymbol{\beta}| < 0$ per 0.3; la soluzione a questa funzione di perdita sarà una stima di $\boldsymbol{\beta}$ tale che la proporzione di scarti $|y_i - \mathbf{x}_i^\top \boldsymbol{\beta}| < 0$ sarà un valore prossimo $\tau = 0.7$, fornendo così una stima della distribuzione di $Y|X$ nel quantile τ . La funzione di perdita definita risulta quindi:

$$\sum_{i: y_i > \mathbf{x}_i^\top \boldsymbol{\beta}} \tau |y_i - \mathbf{x}_i^\top \boldsymbol{\beta}| + \sum_{i: y_i < \mathbf{x}_i^\top \boldsymbol{\beta}} (1 - \tau) |y_i - \mathbf{x}_i^\top \boldsymbol{\beta}|. \quad (1.3)$$

Una scrittura alternativa a (1.3), utilizzata più spesso e che verrà utilizzata anche in seguito, definisce la funzione ρ_τ , detta funzione di perdita quantilica, come la somma degli scarti assoluti pesati per τ :

$$\sum_{i=1}^n \rho_\tau(y_i - \mathbf{x}_i^\top \boldsymbol{\beta}), \quad \text{con} \quad \rho_\tau(z) = \begin{cases} \tau z & \text{se } z \geq 0 \\ (\tau - 1)z & \text{se } z < 0, \end{cases} \quad (1.4)$$

vedasi Hao e Naiman 2007 per ulteriori approfondimenti.

Un'alternativa alla minimizzazione della funzione di perdita (1.4), utilizzabile per ottenere una stima del quantile τ di $Y|X$, è quella della stima tramite massima verosimiglianza. La stima di $\boldsymbol{\beta}$ ottenuta tramite massima verosimiglianza, assumendo che (y_1, \dots, y_n) sia un campione da una distribuzione *Laplace Asimmetrica*, $Y \sim \text{AL}_\tau(\tau, X^\top \boldsymbol{\beta}, \sigma)$, è equivalente a quella che si ottiene minimizzando la funzione di perdita quantilica. La densità *Laplace Asimmetrica*, appartenente, per τ fissato, alla famiglia di distribuzioni di posizione-scala, è una famiglia con densità:

$$\pi_\tau(Y, \mu, \sigma) = \frac{\tau(1 - \tau)}{\sigma} e^{-\rho_\tau\left(\frac{Y - \mu}{\sigma}\right)},$$

con ρ_τ definita in (1.4). Massimizzare una densità *Laplace asimmetrica* in funzione del parametro di posizione μ , con (τ, σ) fissati, equivale a minimizzare l'esponente, quindi a minimizzare la medesima funzione di perdita definita in (1.4). Viene di seguito mostrato anche a livello grafico l'equivalenza tra la funzione di perdita in (1.3) e la densità *Laplace Asimmetrica*.

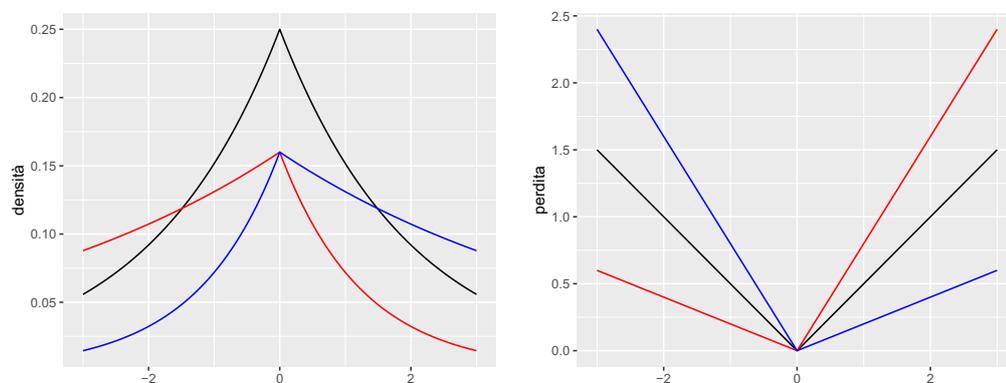


Figura 1.1: Confronto tra densità *Laplace Asimmetrica*, sinistra, e la funzione di perdita ρ_τ , destra, per $\tau = 0.2$ linea blu, $\tau = 0.5$ linea nera, $\tau = 0.8$ linea rossa.

Dato che la distribuzione *Laplace Asimmetrica* è una famiglia di posizione-scala, assumere la seguente distribuzione per Y , $Y \sim \text{AL}(X^\top \boldsymbol{\beta}, \sigma, \tau)$, con densità:

$$\pi(Y|X, \boldsymbol{\beta}, \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} e^{-\rho_\tau\left(\frac{Y-X^\top \boldsymbol{\beta}}{\sigma}\right)}, \quad (1.5)$$

è equivalente ad assumere il modello lineare per Y .

$$Y = X^\top \boldsymbol{\beta} + \varepsilon, \quad (1.6)$$

con $\varepsilon \sim \text{AL}(0, \sigma, \tau)$.

1.2 Regressione quantilica bayesiana

In questa sezione viene brevemente introdotta la modellazione bayesiana, per poi trattare la modellazione quantilica bayesiana univariata e multivariata, introducendo alcuni risultati probabilistici su cui si basano i modelli presentati nel corso del capitolo 3.

1.2.1 Introduzione modellazione bayesiana

La modellazione Bayesiana ha origine nel 1700. con il Teorema di Bayes, e con le elaborazioni successive di Laplace. Nel corso del 1900 la modellazione bayesiana è rimasta a lungo in secondo piano rispetto alla frequentista,

tuttavia negli ultimi decenni, grazie alla potenza di calcolo offerta dai computer, ha trovato sempre maggior utilizzo. Partendo da una distribuzione di probabilità per i parametri di interesse, definita densità a priori, è possibile ricavare una densità, chiamata a posteriori, che tenga conto sia dell'informazione contenuta nella priori, sia dell'informazione sui parametri fornita dalla verosimiglianza. Dato θ vettore dei parametri del modello e \mathbf{y} insieme delle osservazioni, g densità di θ e f densità di \mathbf{y} . si può ottenere la densità a posteriori sfruttando il Teorema di Bayes come:

$$g(\theta|\mathbf{y}) = \frac{g(\theta)f(\mathbf{y}|\theta)}{f(\mathbf{y})} \propto g(\theta)f(\mathbf{y}|\theta).$$

La scelta della distribuzione a priori $g(\theta)$ permette di avere una maggiore flessibilità nella creazione del modello, e può essere usata per veicolare nella stima conoscenze precedenti e assunzioni. Prima dell'avvento dei computer la stima di $g(\theta|\mathbf{y})$ era limitata ad alcuni casi semplici, in cui la struttura di $g(\theta)$ e $f(\mathbf{y}|\theta)$ permettevano la derivazioni di densità note, tuttavia negli anni '90 sono stati presentati algoritmi MCMC, *Markov Chain Monte Carlo*, che hanno reso la modellazione bayesiana possibile anche per modelli estremamente complicati. Questo tipo di algoritmi, che verranno usati anche in seguito, invece di massimizzare direttamente la distribuzione a posteriori del parametro, permettono la generazione da una densità che converge, al crescere delle iterazioni, alla densità a posteriori del parametro; il loro funzionamento verrà spiegato meglio all'interno del capitolo 4.

1.2.2 Modellazione quantilica bayesiana

La prima applicazione di regressione quantilica bayesiana fu proposta da Fatti, Senoana e Thompson 1998, tuttavia furono Yu e Moyeed 2001 i primi a proporre un vero e proprio *framework* bayesiano per la regressione quantilica. Yu e Moyeed 2001, proposero, per stimare il quantile τ della distribuzione $Y|X$, il seguente modello:

$$y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \varepsilon_i,$$

per $i = 1, \dots, n$ con $\varepsilon_i \sim \text{AL}(0, \sigma, \tau)$. La distribuzione a posteriori di $\boldsymbol{\beta}$, parametro di interesse, può essere ricavata, assumendo $\sigma = 1$ come fatto da

Yu e Moyeed 2001, come:

$$\pi(\boldsymbol{\beta}|\mathbf{y}) \propto \pi(\boldsymbol{\beta})\tau^n(1-\tau)^n e^{-\sum_{i=1}^n \rho_\tau(y_i - \mathbf{x}_i^T \boldsymbol{\beta})}.$$

La stima di $\pi(\boldsymbol{\beta})$ può essere fatta tramite algoritmo MCMC. La stima delle distribuzioni dei parametri $(\boldsymbol{\beta}, \sigma)$, pur semplificata dall'utilizzo di algoritmi MCMC, risulta limitata dalla struttura della densità Laplace Asimmetrica, in particolare la non linearità della funzione ρ_τ , che rende impossibile ricondurre la distribuzione a posteriori dei parametri $(\boldsymbol{\beta}, \sigma)$ ad una distribuzione nota. Per poter derivare delle posteriori note, ed utilizzare quindi un algoritmo di tipo *Gibbs Sampling*, Kozumi e Kobayashi 2011, sfruttando i risultati di Kotz, Kozubowski e Podgórski 2001, hanno proposto di utilizzare la scomposizione della distribuzione Laplace asimmetrica come mistura posizione scala di Normali ed Esponenziali per semplificare la derivazione delle densità a posteriori. Kotz, Kozubowski e Podgórski 2001 hanno dimostrato come scomporre in una mistura di posizione scala di gaussiane ed esponenziali $Y \sim \text{AL}(0, \sigma, \tau)$:

$$Y = \lambda v + \varsigma \sqrt{\sigma v} \varepsilon,$$

con $v \sim \text{Exp}(1)$, distribuzione esponenziale di parametro 1, inverso della varianza, e $\varepsilon \sim \text{N}(0, 1)$, normale standard, con $v \perp\!\!\!\perp \varepsilon$. Partendo dalla densità gaussiana condizionata di Y , seguendo Kozumi e Kobayashi 2011, si può dimostrare l'equivalenza tra i modelli in (1.7) e (1.6). La densità condizionata per Y risulta:

$$\pi(Y|\boldsymbol{\beta}, \sigma, \tau) = \frac{1}{\sqrt{2\pi\varsigma^2\sigma v}} \exp\left\{-\frac{(Y - \lambda v)^2}{2\varsigma^2\sigma v}\right\}.$$

La densità di Y marginalizzata rispetto a v si può ottenere integrando $\pi(Y|v)\pi(v)$ in v :

$$\pi(Y|\boldsymbol{\beta}, \sigma, \tau) = \int_0^\infty \frac{1}{\sqrt{2\pi\varsigma^2\sigma v}} \exp\left\{-\frac{(Y - \lambda v)^2}{2\varsigma^2\sigma v}\right\} \frac{1}{\sigma} \exp\left\{-\frac{v}{\sigma}\right\} dv.$$

Per potersi ricondurre alla densità Laplace asimmetrica in (1.5) è necessario definire $\lambda = \frac{1-2\tau}{\tau(1-\tau)}$ e $\varsigma^2 = \frac{2}{\tau(1-\tau)}$ così da centrare il quantile di ordine τ di ε_i in zero. Per approfondimenti e per l'intera dimostrazione si invita a guardare Kozumi e Kobayashi 2011. Il modello in (1.6), contenente anche i termini di regressione per Y dato X , può essere quindi riscritto come mistura:

$$y_i = \mathbf{x}_i^T \boldsymbol{\beta} + \lambda v_i + \varepsilon_i, \tag{1.7}$$

con $\varepsilon_i \sim \mathbf{N}(0, \zeta^2 \sigma v_i)$, $v_i \sim \text{Exp}(\sigma^{-1})$, $\varepsilon_i \perp\!\!\!\perp v_i$ per $i = 1, 2, \dots, n$.

La funzione di verosimiglianza diventa:

$$L(\boldsymbol{\beta}, \sigma, \tau | \mathbf{y}, \mathbf{X}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\zeta^2\sigma v_i}} \exp \left\{ -\frac{(y_i - \mathbf{x}_i^\top \boldsymbol{\beta} - \lambda v_i)^2}{2\zeta^2\sigma v_i} \right\},$$

che si può combinare alla densità a priori di $(\boldsymbol{\beta}, \sigma, v)$, assunte indipendenti, per ottenere la posteriori congiunta $\pi(\boldsymbol{\beta}, \sigma, v | x, y) \propto \pi(\boldsymbol{\beta})\pi(\sigma)\pi(v)L(\boldsymbol{\beta}, \sigma, \tau | y, x)$. In particolare, seguendo Kozumi e Kobayashi 2011, se si assumono come priori $\pi(\boldsymbol{\beta}) \sim N(\boldsymbol{\beta}_\mu, \boldsymbol{\Sigma}_\mu)$, $\pi(\sigma) \sim IG(a, b)$, $\pi(v) \sim \text{Exp}(1)$, la distribuzione a posteriori congiunta diventa:

$$\begin{aligned} \pi(\boldsymbol{\beta}, \sigma, \tau | \mathbf{y}, \mathbf{X}) &= |\boldsymbol{\Sigma}_\beta|^{-1/2} e^{-1/2(\boldsymbol{\beta} - \boldsymbol{\beta}_\mu)^\top \boldsymbol{\Sigma}_\beta (\boldsymbol{\beta} - \boldsymbol{\beta}_\mu)} \\ &\quad \times \sigma^{a+1} e^{-\frac{b}{\sigma}} \\ &\quad \times \prod_{i=1}^n e^{-v_i} \frac{1}{\sqrt{2\pi\zeta^2\sigma v_i}} \exp \left\{ -\frac{(y_i - \mathbf{x}_i^\top \boldsymbol{\beta} - \lambda v_i)^2}{2\zeta^2\sigma v_i} \right\}. \end{aligned}$$

Dalla distribuzione a posteriori congiunta è possibile ricavare delle distribuzioni condizionate per ciascun parametro note:

$$\begin{aligned} \pi(\boldsymbol{\beta} | \mathbf{y}, \mathbf{X}, \sigma, v) &\sim N(\widehat{\boldsymbol{\mu}}_\beta, \widehat{\boldsymbol{\Sigma}}_\beta) \\ \widehat{\boldsymbol{\mu}}_\beta &= \widehat{\boldsymbol{\Sigma}}_\beta \left(\frac{\mathbf{x}_i (y_i - \lambda v_i)}{2\zeta^2\sigma} + \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\mu}_\beta \right) \\ \widehat{\boldsymbol{\Sigma}}_\beta &= \left(\sum_{i=1}^n \left(\frac{\mathbf{x}_i \mathbf{x}_i^\top}{2\zeta^2\sigma} \right) + \boldsymbol{\Sigma}_\beta^{-1} \right)^{-1} \\ \pi(\sigma | \mathbf{y}, \mathbf{X}, \boldsymbol{\beta}, v) &\sim GIG \left(-\alpha - \frac{n}{2}, \frac{\lambda^2 \sum_{i=1}^n v_i}{\zeta^2}, \frac{2b + \sum_{i=1}^n n (y_i - \mathbf{x}_i^\top \boldsymbol{\beta} - \lambda)^2}{\zeta^2 v_i} \right) \\ \pi(v | \mathbf{y}, \mathbf{X}, \sigma, \boldsymbol{\beta}) &\sim GIG \left(\frac{1}{2}, \frac{(y_i - \mathbf{x}_i^\top \boldsymbol{\beta})^2}{\zeta^2 \sigma}, \frac{\lambda^2 \sigma}{\zeta^2} \right), \end{aligned}$$

che possono essere usate per la stima delle distribuzioni tramite algoritmo *MCMC*.

1.2.3 Estensione al caso multidimensionale

Definiamo la densità Laplace asimmetrica d -variata, per $Y \in \mathbb{R}^d$, $Y \sim \text{AL}_d(\boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\Sigma})$, come:

$$\begin{aligned} \pi(Y|\boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\Sigma}) &= \frac{2e^{\boldsymbol{\tau}^\top \boldsymbol{\Sigma}^{-1}(Y-\boldsymbol{\mu})}} (Y-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(Y-\boldsymbol{\mu})}{(2\pi)^{p/2} |\boldsymbol{\Sigma}|^{1/2} (2 + \boldsymbol{\tau}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\tau})} \\ &\quad \times K_{(2-p)/2}(\sqrt{(2 + \boldsymbol{\tau}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\tau})(Y-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(Y-\boldsymbol{\mu})}), \end{aligned}$$

in cui $\boldsymbol{\mu} \in \mathbb{R}^d$ è il parametro di posizione della distribuzione, $\boldsymbol{\Sigma} \in \mathbb{R}_+^{d \times d}$ è la matrice di varianza e $\boldsymbol{\tau} \in \mathbb{R}^d$ è il vettore di quantili in cui centrare la distribuzione, con $K_{(2-p)/2}$ funzione di Bessel del terzo tipo con parametro $(2-p)/2$. La densità Laplace asimmetrica d -variata è anch'essa una distribuzione di tipo posizione-scala, per cui la precedente distribuzione può essere riscritta come:

$$Y = \boldsymbol{\mu} + \boldsymbol{\varepsilon}, \quad (1.8)$$

con $\boldsymbol{\varepsilon} \sim \text{AL}_d(0, \boldsymbol{\tau}, \boldsymbol{\Sigma})$. Siamo interessati anche in questo caso a cercare di riscrivere la densità Laplace asimmetrica multi-variata come mistura di Normali multi-variate. Seguendo Kotz, Kozubowski e Podgórski 2001 si può dimostrare che una densità $Y \sim \text{AL}_d(\boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\Sigma})$ può essere riscritta come mistura di Normali multi-variate come:

$$Y = \boldsymbol{\mu} + \boldsymbol{\lambda} \mathbf{D} v + \boldsymbol{\varsigma} \sqrt{v} \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim \text{N}(0, \boldsymbol{\Sigma}), \quad v \sim \text{Exp}(1),$$

con $\boldsymbol{\lambda}$ e $\boldsymbol{\varsigma}$ vettori in \mathbb{R}^d , $\boldsymbol{\lambda} = (\frac{1-2\tau_1}{\tau_1(1-\tau_1)}, \dots, \frac{1-2\tau_d}{\tau_d(1-\tau_d)})$, $\boldsymbol{\varsigma}^2 = (\frac{2}{\tau_1(1-\tau_1)}, \dots, \frac{2}{\tau_d(1-\tau_d)})$, tali da garantire che $P(Y < \boldsymbol{\mu}) = \tau_j$ per $j = 1, \dots, d$, e \mathbf{D} matrice diagonale contenente le deviazioni standard, ottenuta scomponendo $\boldsymbol{\Sigma}$, come $\boldsymbol{\Sigma} = \mathbf{D} \mathbf{D} \mathbf{D}$, con \mathbf{D} matrice di correlazione. La distribuzione marginale di Y resta $Y \sim \text{AL}_d(\boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\Sigma})$, mentre la distribuzione condizionata, $Y|v$, è una Normale d -variata, $Y|v \sim \text{N}(\boldsymbol{\mu} + \boldsymbol{\lambda} \mathbf{D} v, \boldsymbol{\varsigma}^2 v \boldsymbol{\Sigma})$. Per la dimostrazione fare riferimento a Kotz, Kozubowski e Podgórski 2001.

Riprendendo il modello lineare visto in (1.6), ed estendendolo al caso d -variato, abbiamo:

$$\mathbf{y}_i = \mathbf{B} \mathbf{x}_i + \boldsymbol{\varepsilon}_i,$$

con $\mathbf{B} \in \mathbb{R}^{d \times m}$ matrice di coefficienti di regressione, $\mathbf{x}_i \in \mathbb{R}^m$ vettore contenete le variabili di regressione all'istante i , e $\boldsymbol{\varepsilon}_i \sim \text{AL}_d(0, \boldsymbol{\tau}, \boldsymbol{\Sigma})$. Si può riscrivere il

modello come mistura di esponenziali e normali d -variate come:

$$\mathbf{y}_i = \mathbf{B}\mathbf{x}_i + \boldsymbol{\lambda}\mathbf{D}v_i + \boldsymbol{\varsigma}\sqrt{v_i}\boldsymbol{\varepsilon}_i, \quad \boldsymbol{\varepsilon}_i \sim \mathbf{N}(0, \boldsymbol{\Sigma}), \quad v_i \sim \text{Exp}(1). \quad (1.9)$$

La distribuzione marginale per Y_i , $i = 1, 2, \dots, n$ risulta essere ancora Laplace asimmetrica, $Y_i \sim \text{AL}_d(\mathbf{B}\mathbf{X}_i, \boldsymbol{\tau}, \boldsymbol{\Sigma})$, mentre la distribuzione condizionata, $Y_i|v_i$, diventa una Normale d -variata, $Y_i|v_i \sim N(\mathbf{B}\mathbf{x}_i + \boldsymbol{\lambda}\mathbf{D}v, \boldsymbol{\varsigma}^2v\boldsymbol{\Sigma})$.

Capitolo 2

Il modello autoregressivo vettoriale

In questo capitolo viene brevemente introdotto il modello VAR(p), le sue caratteristiche fondamentali, ed alcune proprietà rilevanti. Viene successivamente presentato il modello VAR con restrizioni sui coefficienti di regressione, che permette una maggiore flessibilità nella definizione del modello. Vengono successivamente presentati alcuni strumenti in grado di aiutare a comprendere meglio le caratteristiche del modello stimato; in particolare vengono presentate *impulse response function* e *forecast error variance decomposition*. Successivamente viene descritto l'approccio bayesiano alla modellazione VAR, vengono introdotte alcune distribuzioni a priori tipicamente usate per i modelli VAR e che verranno riprese in seguito e le priori *spike-and-slab*, una particolare forma gerarchica di distribuzioni a priori che permette di valutare l'importanza dei singoli coefficienti presenti nel modello.

2.1 Il modello VAR(p)

Data una serie ordinata per $i = 1, 2, \dots, n$ di osservazioni multivariate \mathbf{y}_i in \mathbb{R} , il generico modello VAR(p), *Vector Autoregression* di ordine p , viene definito come:

$$\mathbf{y}_i = \boldsymbol{\mu}_y + \sum_{j=1}^p \mathbf{A}_j \mathbf{y}_{i-j} + \boldsymbol{\varepsilon}_i, \quad (2.1)$$

con $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, \dots, \varepsilon_{id})$ in \mathbb{R}^d , vettore delle innovazioni, $E(\boldsymbol{\varepsilon}_i) = 0$, $E(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^\top) = \boldsymbol{\Sigma}_\varepsilon$, $E(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_j) = 0$ per $i \neq j$, $\boldsymbol{\mu}_y = (\mu_{y1}, \mu_{y2}, \dots, \mu_{yd})$ *drift* del modello e $\mathbf{A}_1, \dots, \mathbf{A}_p$ matrici piene di dimensione $d \times d$, con la generica matrice \mathbf{A}_j , $0 < j \leq p$ contenente i coefficienti corrispondenti all'effetto in media dei valori assunti dalla variabile Y_{i-j} sui valori assunti dalla variabile Y_i , con ritardo j rispetto alla variabile precedente.

Il modello può essere riscritto in maniera compatta come:

$$\mathbf{Y} = \mathbf{ZB}^\top + \mathbf{U}, \quad (2.2)$$

con $\mathbf{Z} = (\mathbf{z}_1, \dots, \mathbf{z}_n)^\top$ in $n \times (pd+1)$, $\mathbf{z}_i = (1, \mathbf{y}_{i-1}^\top, \dots, \mathbf{y}_{i-p}^\top)$, $\mathbf{B} = (\boldsymbol{\mu}_y^\top, \mathbf{A}_1^\top, \dots, \mathbf{A}_p^\top)$
 $\mathbf{U} = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \dots, \boldsymbol{\varepsilon}_n)^\top$ in $n \times d$, $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n)^\top$ in $n \times d$.

Qualora si fosse interessati a modellare l'effetto sulla variabile risposta Y di un ulteriore insieme di variabili X , il modello in (2.1) può essere esteso alla forma:

$$\mathbf{y}_i = \boldsymbol{\mu}_y + \sum_{j=1}^p \mathbf{A}_j \mathbf{y}_{i-j} + \mathbf{A}_x \mathbf{x}_i + \boldsymbol{\varepsilon}_i, \quad (2.3)$$

con \mathbf{x}_i vettore di dimensione m di covariate all'istante i e \mathbf{A}_x matrice $d \times m$ la cui l -esima riga rappresenta l'effetto in media di X su Y_l .

Volendo riscrivere il modello in forma compatta come in (2.2) si avrà:

$$\mathbf{Y} = \mathbf{ZB}^\top + \mathbf{U}, \quad (2.4)$$

con $\mathbf{Y} = (\mathbf{z}_1, \dots, \mathbf{z}_n)^\top$ in $n \times (pd + 1 + m)$, $\mathbf{z}_i = (1, \mathbf{y}_{i-1}^\top, \dots, \mathbf{y}_{i-p}^\top, \mathbf{x}_i^\top)$, $\mathbf{B} = (\boldsymbol{\mu}_y^\top, \mathbf{A}_1^\top, \dots, \mathbf{A}_p^\top, \mathbf{A}_x^\top)$, $\mathbf{U} = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \dots, \boldsymbol{\varepsilon}_n)^\top$ in $n \times d$, $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n)^\top$ in $n \times d$.

Per ricondursi ad una forma lineare, stimabile tramite minimi quadrati generalizzati, risulta necessario l'utilizzo dell'operatore vettorizzazione e del prodotto di Kronecker; il modello in (2.4) diventa:

$$\begin{aligned} \mathbf{y} &= (\mathbf{I}_d \otimes \mathbf{Z})\boldsymbol{\beta} + \mathbf{u} \\ \mathbf{y} &= \text{vec}(\mathbf{Y}); \\ \boldsymbol{\beta} &= \text{vec}(\mathbf{B}); \\ \mathbf{u} &= \text{vec}(\mathbf{U}), \mathbf{u} \sim \mathbf{N}_{nd}(0, \boldsymbol{\Sigma}_\varepsilon \otimes \mathbf{I}_n), \end{aligned} \quad (2.5)$$

con $\mathbf{y} \in \mathbb{R}^{nd}$ e $\boldsymbol{\beta} \in \mathbb{R}^{d(1+pd+m)}$.

2.1.1 Stazionarietà

In questa sezione vengono presentate alcune condizioni, stabilità e stazionarietà, fondamentali per la stima e la corretta analisi di un modello VAR. La condizione di stazionarietà per un modello VAR(p) è una condizione necessaria per la stima del modello, per la stima dei suoi momenti e per l'esistenza di una soluzione unica all'equazione di stima; di seguito vengono date le definizioni di stazionarietà e di stabilità per il modello VAR, basandosi sulle definizioni in Lütkepohl 2005.

Definizione: Un modello VAR viene detto debolmente stazionario se $E(Y_i) = E(Y_j)$, $E([Y_i - E(Y_i)]^2) = E([Y_j - E(Y_j)]^2)$, $cov(Y_i, Y_{i+h}) = cov(Y_j, Y_{j+h})$ per ogni i, j in $1, 2, \dots, n$, $h \in \mathbb{N}$, ossia la media e la varianza rimangono invariate in ogni istante $i = 1, 2, \dots, n$ e non dipendono da i .

Definizione: Un modello VAR viene detto stazionario in senso stretto se la distribuzione marginale rimane invariante nel tempo, $f(Y_i, Y_{i+1}, \dots, Y_{i+h}) = f(Y_j, Y_{j+1}, \dots, Y_{j+h})$ $i, j = 1, 2, \dots, n, h \in \mathbb{N}$.

La stazionarietà, sia debole che forte, è una proprietà sui momenti e sulla distribuzione del modello che non risulta facilmente verificabile nella pratica. Un concetto legato alla stazionarietà che può tornare utile è quello della stabilità; infatti la condizione di stabilità su un VAR(p) implica la condizione di stazionarietà. La condizione di stabilità può essere utilizzata quindi per verificare nella pratica, in maniera molto più agevole rispetto all'uso delle definizioni, che un modello VAR stimato per un processo sia stazionario.

Definizione: Un modello VAR(p) è definito stabile se:

$$\det(\mathbf{I}_d - \mathbf{A}_1\chi - \mathbf{A}_2\chi^2 - \dots - \mathbf{A}_p\chi^p) \neq 0, \quad \text{per } |\chi| \leq 1. \quad (2.6)$$

2.2 Il modello autoregressivo vettoriale con restrizioni

Il numero di parametri, al crescere di p , numero di ritardi del modello VAR, d , dimensione della variabile risposta Y , m , numero di covariate da inserire nel modello, cresce in maniera molto rapida come $d(dp+1+m)$. L'elevato numero di parametri presenti in un modello VAR può peggiorare la precisione delle stime, oltre a rendere l'interpretazione del modello maggiormente complicata. Risulta spesso utile o di interesse introdurre delle restrizioni sui parametri del modello, e sulle relazioni tra essi, così da ridurre il numero di parametri da stimare. Queste restrizioni possono essere fornite dalla teoria economica e dal tipo di dati analizzati, ma anche da caratteristiche osservate sulla serie stessa. Se ad esempio si avesse una serie per Y , con d elevato, e i singoli elementi Y_t a memoria lunga, cioè influenzati dai valori a ritardi elevati, un VAR(p) senza restrizioni risulterebbe complicato da stimare, per cui si potrebbe sfruttare un insieme di restrizioni tali da rendere solo i coefficienti autoregressivi per ogni elemento di Y diversi da zero. Di seguito presento il caso di restrizioni lineari sul nostro insieme di parametri di interesse $\boldsymbol{\beta}$. Dato il modello precedentemente introdotto in (2.5):

$$Y = (\mathbf{I}_d \otimes Z)\boldsymbol{\beta} + \mathbf{u}, \quad (2.7)$$

assumiamo un insieme di restrizioni lineari su $\boldsymbol{\beta}$ nella forma:

$$\boldsymbol{\beta} = \text{vec}(\mathbf{B}) = \mathbf{C}\boldsymbol{\delta} + \mathbf{c},$$

con $\mathbf{C} \in (d + d^2p + pm) \times M$ matrice di restrizioni nota, $\boldsymbol{\delta} \in M \times 1$ vettore di parametri di interesse, $\mathbf{c} \in M \times 1$ vettore di costanti note. Assumendo questa forma per $\boldsymbol{\beta}$ il numero di parametri da stimare si riduce da $d + d^2p + pm$ agli M presenti in $\boldsymbol{\delta}$; questa forma è sia in grado di darci flessibilità nel definire il tipo di restrizioni all'interno del modello, sia di ridurre il numero di parametri da stimare tramite specifiche assunzioni. Il nostro modello diventa quindi:

$$\begin{aligned} \mathbf{y} &= (\mathbf{I}_d \otimes \mathbf{Z})(\mathbf{C}\boldsymbol{\delta} + \mathbf{c}) + \mathbf{u} \\ \mathbf{y} - (\mathbf{I}_d \otimes \mathbf{Z})\mathbf{c} &= (\mathbf{I}_d \otimes \mathbf{Z})(\mathbf{C}\boldsymbol{\delta}) + \mathbf{u} \\ \mathbf{y}^* &= (\mathbf{I}_d \otimes \mathbf{Z})(\mathbf{C}\boldsymbol{\delta}) + \mathbf{u}, \end{aligned}$$

con l'ultima forma in $\mathbf{y}^* = \mathbf{y} - (\mathbf{I}_d \otimes \mathbf{Z})\mathbf{c}$ che risulta particolarmente comoda per la stima di $\boldsymbol{\delta}$. Nel caso frequentista si ha infatti che la stima di massima verosimiglianza per $\boldsymbol{\delta}$ si può ottenere come stima ai minimi quadrati per \mathbf{y}^* su $(\mathbf{I}_d \otimes \mathbf{Z})\mathbf{C}$. Per approfondimenti sulla stima in ambito frequentista, e sulle proprietà dello stimatore, fare riferimento a Lütkepohl 2005.

2.3 Analisi Strutturali

I modelli VAR permettono di analizzare e stimare le relazioni che intercorrono tra le variabili di interesse, tuttavia, al crescere di d , numero di variabili risposta modellate, risulta difficile interpretare le relazioni che intercorrono tra di esse semplicemente guardando alle stime dei coefficienti del modello. Per cercare di mitigare questo problema, e fornire alcuni strumenti in grado di aiutare nella comprensione e nella descrizione delle dinamiche del processo VAR di riferimento, vengono presentate due tecniche di analisi strutturale, *impulse response function* e *forecast error variance decomposition*, in grado di aiutare nella comprensione delle relazioni tra variabili all'interno del processo.

2.3.1 Impulse response function

Le *impulse response function* sono uno strumento per valutare l'impatto che uno *shock* al tempo i su una delle variabili di interesse ha sull'intero sistema nel suo sviluppo successivo, in $i + 1, i + 2, \dots$.

Per quantificare l'effetto di uno *shock* iniziale sull'intero sistema si può ricorrere alla rappresentazione VMA del processo VAR; la rappresentazione VMA di un modello VAR è la seguente:

$$\mathbf{y}_i = \boldsymbol{\mu} + \sum_{j=0}^i \boldsymbol{\Phi}_j \boldsymbol{\varepsilon}_{i-j},$$

con $\boldsymbol{\varepsilon}_{i-j} \sim \mathbf{N}_d(0, \boldsymbol{\Sigma}_\varepsilon)$. L'effetto di uno *shock* iniziale sul sistema all'istante i sarà dato dai coefficienti in $\boldsymbol{\Phi}_i$. In particolare, assumendo uno *shock* \mathbf{e}_j , pari a 1 sul j -esimo elemento e 0 altrove, avremo che l'effetto di questo *shock* sul sistema all'istante i verrà dato dalla j -esima riga della matrice $\boldsymbol{\Phi}_i$. La stima delle IRF ottenuta in questo modo non tiene tuttavia conto della dipendenza

delle innovazioni, per cui è adatta solamente nel caso in cui $\boldsymbol{\varepsilon} \sim N_d(0, \boldsymbol{\Sigma}_\varepsilon)$, con $\boldsymbol{\varepsilon}$ a componenti indipendenti.

Viene mostrato di seguito come ricavare le *impulse response function* partendo dal modello VAR(p) in (2.2), senza covariate, in funzione dei parametri.

Dato il modello il forma compatta:

$$\mathbf{y}_i = \mathbf{B}\mathbf{z}_i + \boldsymbol{\varepsilon}_i,$$

con $\boldsymbol{\varepsilon}_i \sim N_d(0, \boldsymbol{\Sigma}_\varepsilon)$, riscrivendo \mathbf{B} e \mathbf{z}_i per evidenziare le componenti autoregressive:

$$\mathbf{y}_i = \boldsymbol{\mu} + \mathbf{A}_1\mathbf{y}_{i-1} + \cdots + \mathbf{A}_p\mathbf{y}_{i-p} + \boldsymbol{\varepsilon}_i,$$

utilizzando la forma canonica del modello si ha:

$$\begin{bmatrix} \mathbf{y}_i \\ \mathbf{y}_{i-1} \\ \mathbf{y}_{i-2} \\ \vdots \\ \mathbf{y}_{i-p+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_{p-1} & \mathbf{A}_p \\ \mathbf{I}_d & 0 & \cdots & 0 & 0 \\ 0 & \mathbf{I}_d & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{I}_d & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y}_{i-1} \\ \mathbf{y}_{i-2} \\ \mathbf{y}_{i-3} \\ \vdots \\ \mathbf{y}_{i-p} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_i \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Il modello in forma canonica può essere considerato come un VAR(1) in \mathbf{y}_i^c :

$$\mathbf{y}_i^c = \begin{bmatrix} \boldsymbol{\mu} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_{p-1} & \mathbf{A}_p \\ \mathbf{I}_d & 0 & \cdots & 0 & 0 \\ 0 & \mathbf{I}_d & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{I}_d & 0 \end{bmatrix} \mathbf{y}_{i-1}^c + \begin{bmatrix} \boldsymbol{\varepsilon}_i \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\mathbf{y}_i^c = \boldsymbol{\mu}_i^c + \mathbf{A}\mathbf{y}_{i-1}^c + \boldsymbol{\varepsilon}_i^c,$$

con $\mathbf{y}_i^c, \boldsymbol{\mu}_i^c, \boldsymbol{\varepsilon}_i^c \in \mathbb{R}^{dp}$. Si ha che, premoltiplicando $\mathbf{J} = [\mathbf{I}_d, 0, \dots, 0]$, matrice in $d \times dp$, per il vettore delle variabili risposta in forma canonica \mathbf{y}_i^c , $\mathbf{J}\mathbf{y}_i^c = \mathbf{y}_i$, si ottiene il modello iniziale per \mathbf{y}_i . Si può riscrivere il modello VAR(1) in forma canonica per \mathbf{y}_i^c nella sua forma in media mobile come:

$$\mathbf{y}_i^c = (\mathbf{I} + \sum_{j=1}^{i-1} \mathbf{A}^j)\boldsymbol{\mu}^c + \mathbf{A}^i\mathbf{y}_0^c + \sum_{j=0}^{i-1} \mathbf{A}^j\boldsymbol{\varepsilon}_{i-j}^c,$$

e

$$\mathbf{y}_i = \mathbf{J}(\mathbf{I} + \sum_{j=1}^{i-1} \mathbf{A}^j) \boldsymbol{\mu}^c + \mathbf{J} \mathbf{A}^i \mathbf{y}_0^c + \sum_{j=0}^{i-1} \mathbf{J} \mathbf{A}^j \mathbf{J}^\top \mathbf{J} \boldsymbol{\varepsilon}_{i-j}.$$

Assumendo $\mathbf{y}_0^c = 0$, e scrivendo con $\boldsymbol{\mu}_i^* = \mathbf{J}(\mathbf{I} + \sum_{j=1}^{i-1} \mathbf{A}^j) \boldsymbol{\mu}^c$, e $\boldsymbol{\Phi}_j = \mathbf{J} \mathbf{A}^j \mathbf{J}^\top$ il modello risulta:

$$\mathbf{y}_i = \boldsymbol{\mu}_i^* + \sum_{j=0}^{i-1} \boldsymbol{\Phi}_j \boldsymbol{\varepsilon}_{i-j}.$$

Ricordando che $\boldsymbol{\varepsilon}_i$ è un vettore di innovazioni, l'effetto di uno *shock* sull'innovazione in 0 sulla variabile \mathbf{y}_j si può ricavare da $\boldsymbol{\Phi}_j$. Il vettore delle innovazioni $\boldsymbol{\varepsilon}_i$, al tempo i , non è a componenti indipendenti, ma ha varianza pari a $\boldsymbol{\Sigma}_\varepsilon$. Per valutare come il sistema risponde ad uno *shock*, su una sola variabile, vorremmo che tale *shock* fosse indotto indipendentemente dai valori delle innovazioni relative alle altre variabili. Una possibile soluzione per avere innovazioni indipendenti è quella di riscrivere il modello come:

$$\begin{aligned} \mathbf{y}_i &= \boldsymbol{\mu}_i^* + \sum_{j=0}^{i-1} \boldsymbol{\Phi}_j \boldsymbol{\Sigma}_\varepsilon^{\frac{1}{2}} \boldsymbol{\Sigma}_\varepsilon^{-\frac{1}{2}} \boldsymbol{\varepsilon}_{i-j} \\ &= \boldsymbol{\mu}_i^* + \sum_{j=0}^{i-1} \boldsymbol{\Phi}_j \boldsymbol{\Sigma}_\varepsilon^{\frac{1}{2}} \boldsymbol{\varepsilon}_{i-j}^0, \end{aligned} \quad (2.8)$$

in cui $\boldsymbol{\Sigma}_\varepsilon = \boldsymbol{\Sigma}_\varepsilon^{\frac{1}{2}} \boldsymbol{\Sigma}_\varepsilon^{\frac{1}{2}\top}$ con $\boldsymbol{\Sigma}_\varepsilon^{\frac{1}{2}}$, matrice triangolare inferiore, decomposizione di Choleski della matrice $\boldsymbol{\Sigma}_\varepsilon$ e con $\boldsymbol{\varepsilon}_{i-j}^0 \sim \mathbf{N}_d(0, \mathbf{I}_d)$, innovazioni indipendenti tra loro. La *impulse response function* al ritardo i per uno *shock* sulla variabile j -esima viene definita come:

$$IRF_i = \mathbf{J} \mathbf{A}^i \mathbf{J}^\top \boldsymbol{\Sigma}_\varepsilon^{\frac{1}{2}} \mathbf{e}_j, \quad (2.9)$$

con \mathbf{e}_j vettore di zeri e uno sulla componente j , che introduce uno *shock* pari alla deviazione standard solo su questa componente. Le IRF così ottenute dipendono dall'ordinamento delle variabili nel modello, sono quindi utilizzate nei casi in cui la teoria economica fornisce già un'ordinamento delle variabili. Per ovviare a questo problema, Pesaran e Shin 1998 hanno proposto l'utilizzo di GIRF, *Generalised impulse response function*, che non dipendano dall'ordinamento, definite come:

$$E(\mathbf{y}_i - \boldsymbol{\mu}^* | \boldsymbol{\varepsilon}_0 = \boldsymbol{\Sigma}_\varepsilon \mathbf{e}_j \sigma_{\varepsilon_{jj}}^{-0.5}) = \mathbf{J} \mathbf{A}^i \mathbf{J}^\top \boldsymbol{\Sigma}_\varepsilon \mathbf{e}_j \sigma_{\varepsilon_{jj}}^{-0.5}, \quad (2.10)$$

che, nel caso con $j = 1$, danno gli stessi risultati delle IRF.

2.3.2 Forecast error variance decomposition

La *forecast error variance decomposition* offre uno spunto diverso rispetto alle *impulse response function*; invece di fornire un'informazione sull'impatto che uno *shock* sulla j -esima variabile ha sulle altre i passi in avanti, la *forecast error variance decomposition* di una variabile l a ritardo i , ci dà un'idea di quanto le innovazioni su ciascuna variabile spieghino della Varianza prevista al tempo i sulla variabile l . Riprendendo il modello nella sua forma VMA, come scritto in (2.8), così da semplificare il calcolo della Varianza spiegata al ritardo i :

$$\mathbf{y}_i = \boldsymbol{\mu}_i^* + \sum_{j=0}^{i-1} \boldsymbol{\Phi}_j \boldsymbol{\Sigma}_\varepsilon^{\frac{1}{2}} \boldsymbol{\varepsilon}_{i-j}^0,$$

con $\boldsymbol{\varepsilon}_{i-j}^0 \sim \mathbf{N}_d(0, \mathbf{I}_d)$. L'errore previsto al tempo i è dunque:

$$\mathbf{y}_i - \boldsymbol{\mu}_i^* = \sum_{j=0}^{i-1} \boldsymbol{\Phi}_j \boldsymbol{\Sigma}_\varepsilon^{\frac{1}{2}} \boldsymbol{\varepsilon}_{i-j}^0.$$

Prendendo a riferimento il l -esimo elemento di \mathbf{y} e riscrivendo per comodità $\boldsymbol{\Theta}_j = \boldsymbol{\Phi}_j \boldsymbol{\Sigma}_\varepsilon^{\frac{1}{2}}$ si ha:

$$\begin{aligned} y_{l,i} - \mu_{l,i}^* &= \sum_{j=0}^{i-1} \boldsymbol{\Theta}_{l1,j} \varepsilon_{1,i-j}^0 + \cdots + \boldsymbol{\Theta}_{ld,j} \varepsilon_{d,i-j}^0 \\ &= \sum_{k=1}^d \boldsymbol{\Theta}_{lk,0} \varepsilon_{k,i}^0 + \cdots + \boldsymbol{\Theta}_{ld,i-1} \varepsilon_{d,1}^0. \end{aligned}$$

A questo punto è possibile ricavare l'errore di previsione per la variabile l , i passi in avanti come:

$$\begin{aligned} E(y_{l,i} - \mu_{l,i}^*)^2 &= \sum_{k=1}^d (\boldsymbol{\Theta}_{lk,0}^2 + \cdots + \boldsymbol{\Theta}_{ld,i-1}^2) \\ &= \sum_{j=0}^{i-1} \sum_{k=1}^d \boldsymbol{\Theta}_{lk,i-1}^2 \\ &= \sum_{j=0}^{i-1} \sum_{k=1}^d (e_l^\top \boldsymbol{\Theta}_j e_k^\top)^2. \end{aligned}$$

La *forecast error variance decomposition* che misura il contributo della variabile k -esima alla Varianza della variabile l -esima i passi in avanti è:

$$FEVD_{lk,i} = \sum_{j=0}^{i-1} (\mathbf{e}_j^\top \Theta_j \mathbf{e}_j)^2 / \sum_{j=0}^{i-1} \sum_{k=1}^d (e_l^\top \Theta_j e_k^\top)^2. \quad (2.11)$$

Anche in questo caso, come visto per le *impulse response function*, vi è un effetto dato dall'ordinamento delle variabili; per cercare di ovviare a questo problema Pesaran e Shin 1998 hanno proposto una versione generalizzata della *forecast error variance decomposition* definita come:

$$GFEVD_{lk,i} = \sum_{j=0}^{i-1} (\mathbf{e}_l^\top \Phi_j \Sigma_\varepsilon \mathbf{e}_k \sigma_{\varepsilon_{kk}}^{-0.5})^2 / \sum_{j=0}^{i-1} \sum_{k=1}^d (e_l^\top \Phi_j \Sigma_\varepsilon \mathbf{e}_k \sigma_{\varepsilon_{kk}}^{-0.5})^2. \quad (2.12)$$

2.4 VAR bayesiano

Riprendo il modello VAR(p) precedentemente introdotto, nella forma vettorizzata, vista in in (2.5), definito come:

$$\mathbf{y} = (\mathbf{I}_d \otimes \mathbf{Z})\boldsymbol{\beta} + \mathbf{u},$$

con \mathbf{Z} contenente intercetta, termini autoregressivi e variabili di regressione, $\mathbf{u} \sim N(0, I_n \otimes \Sigma_\varepsilon)$ vettore delle innovazioni. Passando ad una procedura bayesiana, siamo interessati a conoscere la distribuzione di $\boldsymbol{\beta}$ e Σ_ε condizionatamente ai dati osservati (\mathbf{y}, \mathbf{Z}) . Risulta necessario scegliere una distribuzione a priori per $(\boldsymbol{\beta}, \Sigma_\varepsilon)$ che tenga conto delle caratteristiche di entrambi, in particolare del fatto che Σ_ε è una matrice di covarianza, e come tale è definita positiva e simmetrica. La priori maggiormente usata per $\boldsymbol{\beta}$, nel caso di innovazioni $u \sim N(0, \Sigma_\varepsilon)$, è una distribuzione $N(0, \Sigma_\beta)$. Avendo una verosimiglianza per dati provenienti da una normale si avrà che la distribuzione a posteriori per $\boldsymbol{\beta}$ risulterà nota e corrispondente a $N(\hat{\boldsymbol{\mu}}_\beta^p, \hat{\Sigma}_{\beta^p})$:

$$\begin{aligned} \hat{\boldsymbol{\mu}}_\beta^p &= \hat{\Sigma}_\beta^p ((\mathbf{I}_d \otimes \mathbf{Z})^\top (I_n \otimes \Sigma_\varepsilon)^{-1} \mathbf{y} + \Sigma_\beta^{-1} \boldsymbol{\mu}_\beta) \\ \hat{\Sigma}_\beta^p &= (\Sigma_\beta^{-1} + (\mathbf{I}_d \otimes \mathbf{Z})^\top (I_n \otimes \Sigma_\varepsilon)^{-1} (\mathbf{I}_d \otimes \mathbf{Z}))^{-1}. \end{aligned}$$

In passato sono state proposte varia combinazioni di priori per Σ_ε , con $\boldsymbol{\beta}$ Normale, ad esempio priori Normale-Diffusa e Normale-Wishart; nel seguito

viene presentato il caso Normale-Wishart Inversa, assunte con priori indipendenti tra loro, che verrà anche usato in seguito, poichè risulta essere il più semplice da utilizzare nella derivazione delle densità a posteriori condizionate, utili per eseguire passi di *Gibbs Sampling* all'interno di un algoritmo di stima MCMC. La priori scelta $\Sigma_\varepsilon \sim \text{IW}(c, \mathbf{C})$, corrisponde alla densità:

$$\pi(\Sigma_\varepsilon) = \frac{|\mathbf{C}|^{\frac{c}{2}}}{2^{\frac{dc}{2}} \Gamma_d(\frac{d}{2})} |\Sigma_\varepsilon|^{-\frac{(d+c+1)}{2}} e^{-\frac{\text{tr}(\mathbf{C}\Sigma_\varepsilon^{-1})}{2}},$$

con Γ_p funzione gamma multi-variata.

La distribuzione a posteriori, nel caso $\beta \sim N(0, \Sigma_\beta)$ e $\mathbf{u} \sim N(0, \Sigma_\varepsilon)$, risulta essere $\pi(\Sigma_\varepsilon | \mathbf{y}, \mathbf{X}, \beta) \sim \text{IW}(c^p, \mathbf{C}^p)$, con:

$$\begin{aligned} \mathbf{C}^p &= \mathbf{C} + (\mathbf{Y} - \mathbf{Z}^\top \mathbf{B})^\top (\mathbf{Y} - \mathbf{Z}^\top \mathbf{B}) \\ c^p &= c + n. \end{aligned}$$

2.5 *Spike-and-slab*

Quando il numero di ritardi o di variabili inserite nel modello VAR cresce, può risultare di interesse fare selezione di variabili sulla componente β , in modo da valutare l'importanza delle variabili e selezionare solo le relazioni tra componenti di Y che sembrano avere un reale impatto. In ambito bayesiano uno dei possibili metodi per la selezione di variabili e l'introduzione di sparsità nel modello consiste nell'utilizzo di distribuzioni a priori *Spike-and-slab*, consistenti in una mistura tra una distribuzione con densità fortemente concentrata attorno allo 0, con varianza k_0 , detta *spike*, e una distribuzione con varianza maggiore, detta *slab*, tali che $\text{var}(\text{slab}) \gg k_0$. La probabilità a posteriori, per ciascun parametro, di appartenere alla *spike* o alla *slab*, ci da una stima della significatività o meno dell'effetto della relativa variabile; le variabili con elevata probabilità di appartenere alla *spike* potranno quindi essere considerate come poco influenti nel modello. Di seguito vengono presentate le priori *spike-and-slab* per $\text{vec}(\mathbf{B}) = \beta$ proposte, simili a quelle presenti in Kaufmann e Schumacher 2013, pensate appositamente per un modello VAR.

Le priori *spike-and-slab* proposte introducono sulla matrice diagonale Σ_β ,

varianza di $\boldsymbol{\beta} \sim N(0, \boldsymbol{\Sigma}_\beta)$, una struttura gerarchica, tale che, se l' i -esimo elemento della diagonale di $\boldsymbol{\Sigma}_\beta$ appartiene allo *spike*, avrà una priori fortemente concentrata in zero, schiacciando anche la generazione della posteriori in 0; al contrario se appartiene alla *slab* la priori risulterà poco informativa. Per tenere conto della forma di un modello VAR, contenente d variabili risposta e $M = d + dp + m$ covariate o autoregressori, le *spike-and-slab* di Kaufmann e Schumacher 2013 presentano una struttura che tiene conto, nella varianza di $\boldsymbol{\beta}$, della scala della covariata e del regressore di riferimento all'interno della componente di *spike*. Rispetto a quanto fatto da Kaufmann e Schumacher 2013 ho deciso di optare per una struttura leggermente diversa, in cui la scala della covariata e del regressore non influenzano solo la componente *spike*, ma l'intera varianza di $\boldsymbol{\beta}$, lasciando così il rapporto tra la componente *slab* e *spike* fissato e pari a k_0 . Data B , matrice dei coefficienti del modello, si ha la seguente struttura per il singolo elemento B_{lj} , appartenente alla matrice di coefficienti \mathbf{B} :

$$\begin{aligned} \pi(B_{lj} | \sigma_{\varepsilon_j}, \gamma_{lj}, k_{l1}) &\sim N(0, \sigma_{B_{lj}}) \\ \sigma_{B_{lj}} &= \sigma_{\varepsilon_j} k_{l1} ((1 - \gamma_{lj})k_0 + \gamma_{lj}), \\ 0 \leq k_0 &\ll 1, \quad j = 1, 2, \dots, m \quad l = 1, 2, \dots, d. \end{aligned} \quad (2.13)$$

L'appartenenza per i B_{lj} alla componente *slab* o alla componente *spike* viene regolata dal parametro γ_{lj} , (l, j) -esimo elemento della matrice γ , ottenuto da $\gamma_{lj} \sim \text{Bin}(1, \pi_{0l})$, distribuzione binomiale con una estrazione con probabilità fissata per tutti gli elementi della riga,

$$\pi(\gamma_{lj}) = \pi_{0l}^{\gamma_{lj}} (1 - \pi_{0l})^{1 - \gamma_{lj}}. \quad (2.14)$$

Il parametro k_{l1} misura la variabilità dei \mathbf{B}_l , appartenenti alla componente *slab*, k_{l1} viene assunto con priori Gamma Inversa, $\text{IG}(a_k, b_k)$,

$$\pi(k_{l1}) = \prod_{i=1}^m \frac{k_{i1}^{a_k - 1} (1 - k_{i1})^{-a_k - b_k - 1}}{\text{Be}(a_k + 1, b_k + 1)} \mathbf{I}_{(0, \infty)}(k_{i1}). \quad (2.15)$$

Il parametro σ_{ε_j} , viene assunto con priori Gamma Inversa e viene utilizzato come misura della scala dei $\mathbf{B}_{\cdot j}$

$$\sigma_{\varepsilon_j} \sim \text{IG}(a_\sigma, b_\sigma). \quad (2.16)$$

Il parametro π_{0l} è un parametro che genera la probabilità a priori di generare un elemento appartenente alla *spike* o alla *slab*; il valore di π_{0l} viene dato in input a γ_{lj} ed è lo stesso per ogni riga. Gli iperparametri di π_{0l} sono importanti, in quanto, tramite opportune scelte di (a_π, b_π) è possibile regolare la proporzione di elementi generati dalla *spike*; (a_π, b_π) possono essere quindi considerati congiuntamente come un parametro che introduce maggiore o minore sparsità nel modello. Il parametro π_{0l} viene generato da una distribuzione a priori Beta, $\pi_{0l} \sim \text{Be}(a_\pi, b_\pi)$:

$$\pi(\pi_{0l}) = \frac{\pi_{0l}^{a_\pi-1} (1 - \pi_{0l})^{b_\pi-1}}{\text{Be}(a_\pi, b_\pi)}. \quad (2.17)$$

Le priori *spike-and-slab* qui presentate verranno utilizzate nel corso del capitolo 3, dove verrà anche mostrata la derivazione delle distribuzioni condizionate dato il modello di riferimento.

Capitolo 3

Modello VAR quantilico

L'obiettivo di questo capitolo è definire un modello VARX(p) quantilico bayesiano, presentandone la struttura e descrivendo i possibili algoritmi per la stima di questo. Nella prima parte del capitolo vengono brevemente presentati alcuni articoli sul tema della regressione VAR quantilica, in ambito frequentista. Viene successivamente introdotto e definito un modello VAR quantilico di tipo bayesiano e viene mostrata la derivazione delle distribuzioni a posteriori per i parametri del modello. Viene successivamente presentata l'estensione del modello VAR quantilico al caso di un VAR con restrizioni, in grado di permettere una maggiore flessibilità nella definizione e nella stima del modello. Presento poi una ulteriore estensione, atta a valutare l'importanza dei coefficienti di regressione e fare selezione di variabili e di ritardi, estendendo il modello VAR quantilico tramite l'inclusione di una priori di tipo *spike-and-slab*. Infine, nell'ultima parte di capitolo, introduco una mia proposta per la stima di *impulse response function* e *forecast error variance decomposition* specifiche per un VAR quantilico bayesiano.

3.1 Modelli VAR quantilici in ambito frequentista

In questa sezione sono presentati alcuni articoli sulla modellazione quantilica per modelli VAR in ambito frequentista. Cecchetti e L. Li 2008 , utilizzando la funzione di perdita in (1.4) come funzione di perdita per la stima di un

modello VAR, hanno modellato le dinamiche di prezzi e inflazione. Il contributo più interessante fornito da Cecchetti e L. Li 2008 riguarda la stima della matrice di covarianza per i residui della regressione quantilica sul quantile τ come:

$$\Omega_{\tau} = (\omega_{jk}) = \frac{E[\rho_{\tau}(\varepsilon_{ji})\rho_{\tau}(\varepsilon_{ki})]}{E[f_{\varepsilon}(0)_{ji}]E[f_{\varepsilon}(0)_{ki}]}, \quad j, k \in (1, \dots, d)$$

con ε_{ji} residui del modello VAR sulla variabile j all'istante i . Per ottenere una stima numerica è necessario sostituire a $E[\rho_{\tau}(\varepsilon_{ji})\rho_{\tau}(\varepsilon_{ki})]$ la sua stima empirica:

$$E[\rho_{\tau}(\varepsilon_{ji})\rho_{\tau}(\varepsilon_{ki})] = \frac{1}{n} \sum_{i=1}^n (\tau_j - I(\varepsilon_{ji} < 0))(\tau_k - I(\varepsilon_{ki} < 0)).$$

Per stimare la densità $f_{\varepsilon}(0)_{ki}$ densità del residuo della k -esima variabile in 0, Cecchetti e L. Li 2008 hanno proposto di usare un *Kernel* gaussiano valutato in zero. La matrice di covarianza stimata in questo modo può essere poi utilizzata per ricavare strumenti di analisi del processo stimato quali *impulse response function* e *forecast error variance decomposition*.

Un altro contributo metodologico importante è stato dato da G. Li, Y. Li e Tsai 2015, i quali hanno proposto un metodo per la stima delle autocorrelazioni e delle autocorrelazioni parziali sui quantili. Tramite l'utilizzo delle funzioni di autocorrelazione e autocorrelazione parziale quantilica, G. Li, Y. Li e Tsai 2015 hanno riadattato la procedura di Box-Jenkins per la selezione del modello, proponendo così per prima un procedimento per la selezione del ritardo p all'interno di un modello VAR quantilico.

White, Kim e Manganeli 2015 hanno proposto di utilizzare la stima di un processo VAR in un quantile di riferimento, per poi ricavare una stima del *Value at Risk* relativo al quantile medesimo.

Chavleishvili e Manganeli 2019, partendo dal modello VAR quantilico hanno proposto la stima di un modello VAR strutturale, equazione per equazione, da usare per la previsione di serie di produzione industriale e finanziaria. Un'interessante applicazione, ripresa nel capitolo 5, riguarda l'utilizzazione della modellazione quantilica per fare *stress test*. Questa procedura di *stress testing* viene effettuata assumendo, per un periodo di tempo, realizzazioni del modello in un insieme di quantili estremi, rappresentanti appunto una

situazione di stress, e successivamente facendo ritornare il modello alle osservazioni in mediana, così da valutare l'effetto che la situazione di stress ha creato nel sistema.

3.2 Modello VAR quantilico bayesiano

In ambito bayesiano l'unico modello VAR quantilico è stato proposto da Schüler 2014. In questa sezione viene definito lo stesso modello VAR presentato da Schüler 2014, esteso al caso con covariate VARX, e si mostra la derivazione delle distribuzioni a posteriori, non presente in Schüler 2014, per i parametri di riferimento. Nella derivazione delle posteriori per i singoli parametri, sono state fatte alcune scelte differenti rispetto a quanto fatto da Schüler 2014. Nel seguito di questo paragrafo verrà usata la seguente notazione:

$$\begin{aligned}
\mathbf{Y} &= (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n) \in \mathbb{R}^{d \times n} \\
\mathbf{B} &\in \mathbb{R}^{d \times pd+1+m} \\
\mathbf{z}_i &= [1, \mathbf{y}_{i-1}^\top, \dots, \mathbf{y}_{i-p+1}^\top, \mathbf{x}_i]^\top \in \mathbb{R}^{(dp+1+m) \times 1} \\
\mathbf{Z} &= (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n) \in \mathbb{R}^{(dp+1+m) \times n} \\
\mathbf{V} &= \text{diag}(v_1, v_2, \dots, v_n) \\
\mathbf{y} &= \text{vec}(\mathbf{Y}) \in \mathbb{R}^{dn \times 1} \\
\boldsymbol{\beta} &= \text{vec}(\mathbf{B}) \in \mathbb{R}^{(d^2p+d+dm) \times 1} \\
\mathbf{v} &= (v_1, v_2, \dots, v_n) \\
\mathbf{S}_\tau &= \text{diag}(\boldsymbol{\varsigma}) \\
\boldsymbol{\Sigma} &= \mathbf{v} \otimes \mathbf{D} \mathbf{S}_\tau \mathbf{R} \mathbf{S}_\tau \mathbf{D}.
\end{aligned}$$

Definiamo il seguente modello sul vettore di quantili di interesse $\boldsymbol{\tau} = (\tau_1, \dots, \tau_d)$ con $\mathbf{y}_i \in \mathbb{R}^d$ vettore di variabili di interesse e $\mathbf{z}_i \in \mathbb{R}^{dp+d+m}$ contenente i termini autoregressivi fino a ritardo p , drift e covariate, per $i = 1, \dots, n$.

$$\begin{aligned}
\mathbf{y}_i &= \mathbf{B} \mathbf{z}_i + \boldsymbol{\varepsilon}_i \\
\mathbf{y}_i &= \boldsymbol{\mu} + \mathbf{A}_1 \mathbf{y}_{i-1} + \dots + \mathbf{A}_p \mathbf{y}_{i-p} + \mathbf{A}_x \mathbf{x}_i + \boldsymbol{\varepsilon}_i,
\end{aligned} \tag{3.1}$$

con $\varepsilon_i \sim AL_d(0, \mathbf{DRD}, \boldsymbol{\tau})$. Questo equivale ad assumere la seguente distribuzione condizionata per \mathbf{y}_i :

$$\mathbf{y}_i | \mathbf{B}, \mathbf{z}_i \sim AL_d(\mathbf{Bz}_i, \mathbf{DRD}, \boldsymbol{\tau}). \quad (3.2)$$

Il modello è scrivibile alternativamente, sfruttando le proprietà presentate nella sezione 1.2 come mistura di Normali multivariate, grazie all'introduzione di fattori latenti \mathbf{v} con distribuzione esponenziale. Il modello può quindi essere riscritto equivalentemente come:

$$\mathbf{y}_i = \mathbf{Bz}_i + v_i \mathbf{D}\boldsymbol{\lambda} + \sqrt{v_i} \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}} \mathbf{u}_i,$$

con $\mathbf{u}_i \sim N_d(0, \mathbf{I}_d)$ e $\boldsymbol{\Omega}^{\frac{1}{2}}$, decomposizione di Choleski di $\boldsymbol{\Omega}$ tale che $\boldsymbol{\Omega}^{\frac{1}{2}} \boldsymbol{\Omega}^{\frac{1}{2}\top} = \boldsymbol{\Omega}$, con $\boldsymbol{\lambda} = (\frac{1-2\tau_1}{\tau_1(1-\tau_1)}, \dots, \frac{1-2\tau_d}{\tau_d(1-\tau_d)})$ e $\boldsymbol{\varsigma}^2 = (\frac{2}{\tau_1(1-\tau_1)}, \dots, \frac{2}{\tau_d(1-\tau_d)})$, per $\boldsymbol{\tau} = (\tau_1, \dots, \tau_d)$ vettore dei quantili di interesse. La matrice $\boldsymbol{\Omega}$ viene ottenuta come $\boldsymbol{\Omega} = \mathbf{S}_\tau \mathbf{R} \mathbf{S}_\tau$, con \mathbf{R} matrice di correlazione, e $\mathbf{S}_\tau = \text{diag}(\boldsymbol{\varsigma})$. La matrice $\mathbf{D} = \text{diag}(d_{11}, \dots, d_{dd})$ è semplicemente una matrice diagonale contenente le deviazioni standard del modello. I fattori latenti $v_i \sim \text{Exp}(1)$ indipendenti ed identicamente distribuiti per $i = 1, 2, \dots, n$. Il modello può essere scritto, alternativamente, sfruttando l'operatore vettorizzazione, come:

$$\mathbf{y} = \text{vec}(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{v} \otimes (\mathbf{D}\boldsymbol{\lambda}) + (\sqrt{\mathbf{v}} \otimes (\mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}})) \mathbf{u}, \quad (3.3)$$

con $\mathbf{u} \sim N_{dn}(0, \mathbf{I}_{dn})$. La distribuzione per \mathbf{y}_i condizionatamente ai fattori latenti risulta essere:

$$\mathbf{y}_i | v_i \mathbf{X}_i, \mathbf{B}, \boldsymbol{\Omega}, \mathbf{D} \sim N_d(\mathbf{Bz}_i + v_i \boldsymbol{\lambda}^\top \mathbf{D}, v_i \mathbf{D}\boldsymbol{\Omega}\mathbf{D}).$$

I parametri di interesse sono $(\boldsymbol{\beta}, \mathbf{D}, \boldsymbol{\Omega})$ ed è necessario conoscere la distribuzione di \mathbf{v} , vettore dei fattori latenti, a posteriori. Le priori proposte da Schüler 2014, e che vengono riprese in questa prima parte, sono le seguenti:

$$\pi(v_i) = e^{-v_i} \quad (3.4)$$

$$\pi(\boldsymbol{\beta}) = |2\pi \boldsymbol{\Sigma}_\beta|^{-\frac{1}{2}} e^{-(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)^\top (2\boldsymbol{\Sigma}_\beta)^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)} \quad (3.5)$$

$$\pi(\boldsymbol{\Omega}) \propto |\mathbf{C}_\Omega|^{-(c-d-1)} e^{-\frac{\text{Tr}(\mathbf{C}_\Omega \boldsymbol{\Omega}^{-1})}{2}} \quad (3.6)$$

$$\pi(\mathbf{D}) \propto 1. \quad (3.7)$$

Date le priori, la distribuzione congiunta del modello risulta la seguente:

$$\begin{aligned} \pi(\mathbf{Y}, \boldsymbol{\beta}, \mathbf{D}, \boldsymbol{\Omega}, \mathbf{v} | \mathbf{X}) &\propto |\boldsymbol{\Sigma}|^{-1/2} \\ &\times e^{-\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})^\top (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})} \\ &\times e^{-\sum_1^n v_i |\mathbf{C}\boldsymbol{\Omega}|^{-(c-d-1)} e^{-\frac{\text{tr}(\mathbf{C}\boldsymbol{\Omega}\boldsymbol{\Omega}^{-1})}{2}}}. \end{aligned}$$

Di seguito vengono ricavate le distribuzioni a posteriori per i singoli parametri $(\mathbf{v}, \mathbf{B}, \boldsymbol{\Omega}, \mathbf{D})$, condizionatamente agli altri, che verranno successivamente utilizzate per la stima del modello tramite algoritmo MCMC. La distribuzione a posteriori condizionata di $\boldsymbol{\beta}$ risulta proporzionale a:

$$\begin{aligned} \pi(\boldsymbol{\beta} | \mathbf{D}, \boldsymbol{\Omega}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})^\top (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})} \\ &\times e^{-\frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)^\top \boldsymbol{\Sigma}_\beta^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)}. \end{aligned}$$

Vengono isolati gli elementi che dipendono da $\boldsymbol{\beta}$ ad esponente:

$$\begin{aligned} \pi(\boldsymbol{\beta} | \boldsymbol{\Omega}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\frac{1}{2}\boldsymbol{\beta}^\top \mathbf{X}^\top (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} \mathbf{X}\boldsymbol{\beta}} \\ &\times e^{\frac{1}{2}\boldsymbol{\beta}^\top \mathbf{X}^\top (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{y} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})} \\ &\times e^{-\frac{1}{2}\boldsymbol{\beta}^\top \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\beta}} e^{\boldsymbol{\beta}^\top \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\mu}_\beta}. \end{aligned}$$

Raggruppando gli esponenti si ottiene:

$$\begin{aligned} \pi(\boldsymbol{\beta} | \boldsymbol{\Omega}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\frac{1}{2}\boldsymbol{\beta}^\top (\mathbf{X}^\top (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} \mathbf{X} + \boldsymbol{\Sigma}_\beta^{-1}) \boldsymbol{\beta}} \\ &\times e^{\boldsymbol{\beta}^\top (\mathbf{X}^\top (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{y} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda}) + \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\mu}_\beta)}, \end{aligned}$$

di conseguenza la posteriori di $\boldsymbol{\beta}$ corrisponde a $\mathcal{N}(\hat{\boldsymbol{\mu}}_\beta^p, \hat{\boldsymbol{\Sigma}}_\beta^p)$:

$$\begin{aligned} \hat{\boldsymbol{\mu}}_\beta^p &= \hat{\boldsymbol{\Sigma}}_\beta^p (\mathbf{X}^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda}) + \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\mu}_\beta) \\ \hat{\boldsymbol{\Sigma}}_\beta^p &= (\boldsymbol{\Sigma}_\beta^{-1} + \mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}. \end{aligned} \tag{3.8}$$

Ricordando la scrittura come prodotto di Kronecker e tramite l'operatore vettorizzazione, si può riscrivere il tutto come:

$$\begin{aligned} \hat{\boldsymbol{\mu}}_\beta^p &= \hat{\boldsymbol{\Sigma}}_\beta^p ((\mathbf{Z}^\top \otimes \mathbf{I}_d) (\mathbf{V}^{-1} \otimes (\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1}) \text{vec}((\mathbf{Y} - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})^\top) + \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\mu}_\beta) \\ \hat{\boldsymbol{\Sigma}}_\beta^p &= (\boldsymbol{\Sigma}_\beta^{-1} + (\mathbf{Z} \otimes \mathbf{I}_d) (\mathbf{V}^{-1} \otimes (\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1}) (\mathbf{Z}^\top \otimes \mathbf{I}_d))^{-1}. \end{aligned}$$

Sfruttando le proprietà del prodotto di Kronecker e dell'operatore vettorizzazione si ottiene la seguente forma:

$$\begin{aligned} \hat{\boldsymbol{\mu}}_\beta^p &= \hat{\boldsymbol{\Sigma}}_\beta^p (\text{vec}((\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{Y} - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})^\top \mathbf{V}^{-1} \mathbf{Z}^\top) + \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\mu}_\beta) \\ \hat{\boldsymbol{\Sigma}}_\beta^p &= (\boldsymbol{\Sigma}_\beta^{-1} + (\mathbf{Z}\mathbf{V}^{-1}\mathbf{Z}^\top) \otimes (\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1})^{-1}. \end{aligned}$$

La forma così ottenuta offre, rispetto alla forma in (3.8), un vantaggio computazionale dato dal non dover operare su matrici ad elevata dimensionalità, frutto di prodotti di Kronecker, ma di operare su matrici di minore dimensione e di combinare successivamente i risultati tramite prodotto di Kronecker. Questo vantaggio diventa sempre più importante al crescere di p e n . In alternativa alla priori informativa proposta da Schüler 2014, propongo, per semplificare e velocizzare ulteriormente la stima dei parametri a posteriori della distribuzione di β , di assumere una priori impropria completamente non informativa, $\pi(\beta) \propto 1$. In questo caso la densità congiunta del modello risulta:

$$\begin{aligned} \pi(\beta, \Omega, \mathbf{D}, \mathbf{v} | \mathbf{Y}, \mathbf{X}) &\propto |\Sigma|^{-1/2} e^{-\frac{1}{2}(\mathbf{y} - \mathbf{X}\beta - \mathbf{v} \otimes \mathbf{D}\lambda)^\top (\Sigma)^{-1} (\mathbf{y} - \mathbf{X}\beta - \mathbf{v} \otimes \mathbf{D}\lambda)} \\ &\times e^{-\sum_{i=1}^n v_i} |\mathbf{C}_\Omega|^{-(c-d-1)} e^{-\frac{\text{Tr}(\mathbf{C}_\Omega \Omega^{-1})}{2}}. \end{aligned}$$

La distribuzione a posteriori per β risulta proporzionale a:

$$\begin{aligned} \pi(\beta | \Omega, \mathbf{D}\mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\frac{1}{2}(\mathbf{y} - \mathbf{X}\beta - \mathbf{v} \otimes \mathbf{D}\lambda)^\top (\Sigma)^{-1} (\mathbf{y} - \mathbf{X}\beta - \mathbf{v} \otimes \mathbf{D}\lambda)} \\ &\propto e^{-\frac{1}{2}(\mathbf{X}\beta)^\top (\Sigma)^{-1} (\mathbf{X}\beta)} e^{(\mathbf{X}\beta)^\top (\Sigma)^{-1} (\mathbf{y} - \mathbf{v} \otimes \mathbf{D}\lambda)}, \end{aligned}$$

di conseguenza la posteriori di β corrisponde a $\mathbf{N}(\hat{\mu}_\beta^p, \hat{\Sigma}_\beta^p)$:

$$\begin{aligned} \hat{\mu}_\beta^p &= \hat{\Sigma}_\beta^p ((\mathbf{Z}^\top \otimes \mathbf{I}_d)(\mathbf{V}^{-1} \otimes (\mathbf{D}\Omega\mathbf{D})^{-1}) \text{vec}((\mathbf{Y} - \mathbf{v} \otimes \lambda^\top \mathbf{D})^\top)) \\ \hat{\Sigma}_\beta^p &= ((\mathbf{Z} \otimes \mathbf{I}_d)(\mathbf{V}^{-1} \otimes (\mathbf{D}\Omega\mathbf{D})^{-1})(\mathbf{Z}^\top \otimes \mathbf{I}_d))^{-1}. \end{aligned}$$

Sfruttando le proprietà del prodotto di Kronecker e dell'operatore vettorizzazione si ottiene:

$$\begin{aligned} \hat{\mu}_\beta^p &= \text{vec}((\mathbf{Y} - \mathbf{v} \otimes \lambda^\top \mathbf{D})^\top \mathbf{V}^{-1} \mathbf{Z}^\top (\mathbf{Z} \mathbf{V}^{-1} \mathbf{Z}^\top)^{-1}) \\ \hat{\Sigma}_\beta^p &= ((\mathbf{Z} \mathbf{V}^{-1} \mathbf{Z}^\top)^{-1}) \otimes \mathbf{D}\Omega\mathbf{D}. \end{aligned} \tag{3.9}$$

La forma così ottenuta ha il vantaggio di non dover più invertire il prodotto di Kronecker a dimensione elevata in $\hat{\Sigma}_\beta^p$, ma di operare prima l'inversione in dimensione minore e poi il prodotto di Kronecker; la stima di $\hat{\mu}_\beta^p$ invece non dipende più direttamente da Ω . Per quanto riguarda \mathbf{V} , matrice diagonale contenente le realizzazioni del fattore latente \mathbf{v} , la posteriori condizionata risulta proporzionale a:

$$\begin{aligned} \pi(\mathbf{V} | \beta, \mathbf{D}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto |\mathbf{V}|^{-\frac{d}{2}} \\ &\times e^{-\frac{1}{2}(\mathbf{y} - \mathbf{X}\beta - \mathbf{v} \otimes \mathbf{D}\lambda)^\top (\mathbf{V} \otimes \mathbf{D}\Omega\mathbf{D})^{-1} (\mathbf{y} - \mathbf{X}\beta - \mathbf{v} \otimes \mathbf{D}\lambda)} e^{-\sum_{i=1}^n v_i}. \end{aligned}$$

Sfruttando il fatto che \mathbf{V} è una matrice diagonale, possiamo riscrivere il modello come:

$$\begin{aligned}
\pi(\mathbf{V}|\boldsymbol{\beta}, \mathbf{D}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\sum_{i=1}^n \frac{1}{2}(\mathbf{y}_i - \mathbf{Z}_i \mathbf{B} - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})^\top (v_i \boldsymbol{\Omega})^{-1} (\mathbf{y}_i - \mathbf{Z}_i \mathbf{B} - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})} \\
&\quad \times \prod_{i=1}^n v_i^{-\frac{d}{2}} e^{-v_i} \\
\pi(\mathbf{V}|\boldsymbol{\beta}, \mathbf{D}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\sum_{i=1}^n \frac{1}{2}(\mathbf{y}_i - \mathbf{Z}_i \mathbf{B})^\top (v_i \mathbf{D} \boldsymbol{\Omega} \mathbf{D})^{-1} (\mathbf{y}_i - \mathbf{Z}_i \mathbf{B})} \\
&\quad \times e^{-\sum_{i=1}^n \frac{1}{2}(v_i \mathbf{D} \boldsymbol{\lambda})^\top (v_i \mathbf{D} \boldsymbol{\Omega} \mathbf{D})^{-1} (v_i \mathbf{D} \boldsymbol{\lambda})} \\
&\quad \times e^{-\sum_{i=1}^n \frac{1}{2}(\mathbf{y}_i - \mathbf{Z}_i \mathbf{B})^\top (v_i \mathbf{D} \boldsymbol{\Omega} \mathbf{D})^{-1} (\mathbf{y}_i - \mathbf{Z}_i \mathbf{B} - \boldsymbol{\lambda} v_i \sqrt{\mathbf{D}^\top})} \\
&\quad \times \prod_{i=1}^n v_i^{-\frac{d}{2}} e^{-v_i}.
\end{aligned}$$

A questo punto si possono eliminare le parti esponenziali non contenenti \mathbf{v}

$$\begin{aligned}
\pi(\mathbf{V}|\boldsymbol{\beta}, \mathbf{D}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\sum_{i=1}^n \frac{1}{2v_i} (\mathbf{y}_i - \mathbf{Z}_i \mathbf{B})^\top (\mathbf{D} \boldsymbol{\Omega} \mathbf{D})^{-1} (\mathbf{y}_i - \mathbf{Z}_i \mathbf{B})} \\
&\quad \times e^{-\sum_{i=1}^n \frac{v_i}{2} (\mathbf{D} \boldsymbol{\lambda})^\top (\mathbf{D} \boldsymbol{\Omega} \mathbf{D})^{-1} (\mathbf{D} \boldsymbol{\lambda})} \\
&\quad \times \prod_{i=1}^n v_i^{\frac{(2-d)}{2} + 1} e^{-\frac{2v_i}{2}} \\
\pi(\mathbf{V}|\boldsymbol{\beta}, \mathbf{D}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto \prod_{i=1}^n v_i^{\frac{(2-d)}{2} + 1} e^{-\frac{1}{2v_i} (\mathbf{y}_i - \mathbf{Z}_i \mathbf{B})^\top (\mathbf{D} \boldsymbol{\Omega} \mathbf{D})^{-1} (\mathbf{y}_i - \mathbf{Z}_i \mathbf{B})} \\
&\quad \times e^{-\frac{v_i}{2} (2 + \boldsymbol{\lambda}^\top \mathbf{D} (\mathbf{D} \boldsymbol{\Omega} \mathbf{D})^{-1} (\mathbf{D} \boldsymbol{\lambda}))}.
\end{aligned}$$

La distribuzione condizionata risulta essere una Gaussiana inversa generalizzata, $v_i \sim \text{Gig}(\frac{2-d}{2}, 2 + \boldsymbol{\lambda}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{\lambda}, \boldsymbol{\varepsilon}_i^\top (\mathbf{D} \boldsymbol{\Omega} \mathbf{D})^{-1} \boldsymbol{\varepsilon}_i)$, con $\boldsymbol{\varepsilon}_i = \mathbf{y}_i - \mathbf{Z}_i \mathbf{B}$ vettore dei residui di regressione all'istante i .

Per quanto riguarda la matrice $\boldsymbol{\Omega} = \mathbf{S}_\tau \mathbf{R} \mathbf{S}_\tau$, contenente i termini \mathbf{S}_τ noti e la matrice di correlazione \mathbf{R} , la distribuzione a posteriori condizionata risulta proporzionale a:

$$\begin{aligned}
\pi(\boldsymbol{\Omega}|\mathbf{v}, \boldsymbol{\beta}, \mathbf{D}, \mathbf{Y}, \mathbf{X}) &\propto |\boldsymbol{\Omega}|^{-t/2} \\
&\quad \times e^{-\frac{1}{2}(\mathbf{y} - \mathbf{X} \boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D} \boldsymbol{\lambda})^\top \mathbf{V} \otimes \boldsymbol{\Omega}^{-1} (\mathbf{y} - \mathbf{X} \boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D} \boldsymbol{\lambda})} \\
&\quad \times |\boldsymbol{\Omega}|^{-\frac{c+p+1}{2}} e^{-\frac{\text{Tr}(\mathbf{C} \boldsymbol{\Omega}^{-1})}{2}}.
\end{aligned}$$

Usando l'operatore traccia sulla prima parte dell'equazione si ottiene:

$$\begin{aligned} \pi(\boldsymbol{\Omega}|\boldsymbol{v}, \boldsymbol{\beta}, \mathbf{D}, \mathbf{Y}, \mathbf{X}) &\propto |\boldsymbol{\Omega}|^{-n/2} \\ &\times e^{-\frac{1}{2}\text{tr}(\mathbf{D}^{-1}(\mathbf{Y}-\mathbf{Z}^T\boldsymbol{\beta}-\boldsymbol{v}\otimes\boldsymbol{\lambda}^T\mathbf{D})^T\mathbf{V}^{-1}(\mathbf{Y}-\mathbf{Z}^T\boldsymbol{\beta}-\boldsymbol{v}\otimes\boldsymbol{\lambda}^T\mathbf{D})\mathbf{D}^{-1}\boldsymbol{\Omega}^{-1})} \\ &\times |\boldsymbol{\Omega}|^{-\frac{c+p+1}{2}} e^{-\frac{\text{tr}(\mathbf{C}_{\boldsymbol{\Omega}}\boldsymbol{\Omega}^{-1})}{2}}. \end{aligned}$$

Raccogliendo in $\boldsymbol{\Omega}$ si ha:

$$\begin{aligned} \pi(\boldsymbol{\Omega}|\boldsymbol{v}, \boldsymbol{\beta}, \mathbf{D}, \mathbf{Y}, \mathbf{X}) &\propto |\boldsymbol{\Omega}|^{-\frac{c+p+1+n}{2}} \\ &\times e^{-\frac{\text{tr}((\mathbf{C}_{\boldsymbol{\Omega}}+\mathbf{D}^{-1}(\mathbf{Y}-\mathbf{Z}^T\boldsymbol{\beta}-\boldsymbol{v}\otimes\boldsymbol{\lambda}^T\mathbf{D})^T\mathbf{V}^{-1}(\mathbf{Y}-\mathbf{Z}^T\boldsymbol{\beta}-\boldsymbol{v}\otimes\boldsymbol{\lambda}^T\mathbf{D})\mathbf{D}^{-1})\boldsymbol{\Omega}^{-1})}{2}}. \end{aligned}$$

La distribuzione a posteriori condizionata di $\boldsymbol{\Omega}$ risulta quindi essere:

$\boldsymbol{\Omega}|\boldsymbol{v}, \boldsymbol{\beta}, \mathbf{D}, \mathbf{Y}, \mathbf{X} \sim \text{IW}(\widehat{\mathbf{C}}_{\boldsymbol{\Omega}}^p, \widehat{c}^p)$, con:

$$\begin{aligned} \widehat{\mathbf{C}}_{\boldsymbol{\Omega}}^p &= \mathbf{C}_{\boldsymbol{\Omega}} + \mathbf{D}^{-1}(\mathbf{Y} - \mathbf{Z}^T\boldsymbol{\beta} - \boldsymbol{v} \otimes \boldsymbol{\lambda}^T\mathbf{D})^T\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{Z}^T\boldsymbol{\beta} - \boldsymbol{v} \otimes \boldsymbol{\lambda}^T\mathbf{D})\mathbf{D}^{-1} \\ \widehat{c}^p &= c + n. \end{aligned} \tag{3.10}$$

Infine, per quanto riguarda la stima di \mathbf{D} , ho deciso di procedere diversamente da quanto fatto da Schüler 2014. Schüler 2014 ha ricavato la distribuzione di \mathbf{D} , condizionata a $(\boldsymbol{\Omega}, \boldsymbol{\beta})$, trattando quindi la verosimiglianza di y marginalmente come Laplace asimmetrica, definendo per \mathbf{D} la posteriori come:

$$\begin{aligned} \pi(\mathbf{D}|\mathbf{Y}, \mathbf{X}, \boldsymbol{\Omega}, \boldsymbol{\beta}) &\propto \prod_{i=1}^n \frac{2\exp((\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta})^T(\mathbf{D}\boldsymbol{\Omega})^{-1}\boldsymbol{\lambda})}{(2\pi)^{d/2}|\mathbf{D}\boldsymbol{\Omega}\mathbf{D}|^{1/2}} \\ &\times \left(\frac{(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta})^T(\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta})}{2 + \boldsymbol{\lambda}^T\boldsymbol{\Omega}\boldsymbol{\lambda}} \right)^{-d/2+1} \\ &\times K_{-d/2+1}(\sqrt{(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta})^T(\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta})(2 + \boldsymbol{\lambda}^T\boldsymbol{\Omega}\boldsymbol{\lambda})}). \end{aligned}$$

Data la forma alquanto complicata di $\pi(\mathbf{D}|\mathbf{Y}, \mathbf{X}, \boldsymbol{\Omega}, \boldsymbol{\beta})$, ho deciso di ricavare la posteriori di \mathbf{D} condizionatamente a $(\boldsymbol{\Omega}, \boldsymbol{\beta}, \boldsymbol{v})$, trattando quindi la verosimiglianza normale di $y|\boldsymbol{v}$. La distribuzione a posteriori per \mathbf{D} in questo caso risulta proporzionale a:

$$\begin{aligned} \pi(\mathbf{D}|\boldsymbol{v}, \boldsymbol{\beta}, \boldsymbol{\Omega}, \mathbf{Y}, \mathbf{X}) &\propto |\mathbf{D}|^{-n} \\ &\times e^{-\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\boldsymbol{v}\otimes\mathbf{D}\boldsymbol{\lambda})^T(\mathbf{V}\otimes\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\boldsymbol{v}\otimes\mathbf{D}\boldsymbol{\lambda})} \\ &\propto |\mathbf{D}|^{-n} e^{-\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{V}\otimes\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})} \\ &\times e^{(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{1}_n\otimes(\mathbf{D}\boldsymbol{\Omega})^{-1}\boldsymbol{\lambda})}, \end{aligned}$$

che ha anche il vantaggio di poter essere scritta con gli elementi in forma matriciale. La criticità nel trattamento della posteriori condizionata è data dalla presenza del termine \mathbf{D} sia nella media condizionata di \mathbf{Y} , con il termine $\mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda}$, sia nella varianza di \mathbf{Y} , $\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D}$. La posteriori di \mathbf{D} non risulta riconducibile a nessuna distribuzione nota.

3.2.1 Generazione unica della matrice di covarianza

Nelle sezioni precedenti si è lavorato con la matrice di covarianza per l' i -esimo elemento scomposta in \mathbf{D} , matrice diagonale contenente la deviazione standard e $\boldsymbol{\Omega} = \mathbf{S}_\tau \mathbf{R} \mathbf{S}_\tau$, trattandole separatamente. Nel caso in cui d sia di dimensione ridotte, quindi con un numero di elementi all'interno di $\boldsymbol{\Omega}$ e \mathbf{D} abbastanza limitato, potrebbe risultare una alternativa valida trattare direttamente $\boldsymbol{\Omega}_d = \mathbf{D} \mathbf{S}_\tau \mathbf{R} \mathbf{S}_\tau \mathbf{D}$ invece che di \mathbf{D} e $\boldsymbol{\Omega}$ separatamente. In questo caso, al posto di stimare due distribuzioni a posteriori per i parametri $\boldsymbol{\Omega}$ e \mathbf{D} , propongo di trattarli congiuntamente all'interno della matrice $\boldsymbol{\Omega}_d = \mathbf{D} \mathbf{S}_\tau \mathbf{R} \mathbf{S}_\tau \mathbf{D}$ e calcolare una unica distribuzione a posteriori su cui applicare poi un unico *step Metropolis Hasting*. Di seguito presento i calcoli per calcolare la densità a posteriori condizionata di $\boldsymbol{\Omega}_d$. Come distribuzione a priori per $\boldsymbol{\Omega}_d$ propongo di assumere una distribuzione impropria $\pi(\boldsymbol{\Omega}_d) \sim 1$. Si ha la seguente distribuzione condizionata per \mathbf{y}_i :

$$\pi(\mathbf{y}_i | \mathbf{X}_i, \mathbf{B}, \boldsymbol{\Omega}_d, v_i) = \phi_d(\mathbf{X}_i \mathbf{B} + \mathbf{D} \boldsymbol{\lambda} v_i, v_i \mathbf{S}_\tau \boldsymbol{\Omega}_d \mathbf{S}_\tau). \quad (3.11)$$

La posteriori risulta proporzionale a:

$$\begin{aligned} \pi(\boldsymbol{\Omega}_d | \mathbf{v}, \boldsymbol{\beta}, \mathbf{Y}, \mathbf{X}) &\propto |\boldsymbol{\Omega}_d|^{-t/2} \\ &\times e^{-\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})^\top \mathbf{v} \otimes \mathbf{S}_\tau \boldsymbol{\Omega}_d \mathbf{S}_\tau^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})} \\ &\propto |\boldsymbol{\Omega}_d|^{-t/2} e^{-\frac{1}{2} \text{Tr}(\mathbf{V}^{-\frac{1}{2}} (\mathbf{Y} - \mathbf{Z}\mathbf{B} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda}) \mathbf{S}_\tau^{-1} (\boldsymbol{\Omega}_d)^{-1} \mathbf{S}_\tau^{-1} (\mathbf{Y} - \mathbf{Z}\mathbf{B} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})^\top \mathbf{V}^{-\frac{1}{2}})} \\ &\propto |\boldsymbol{\Omega}_d|^{-\frac{n}{2}} e^{-\frac{1}{2} \text{Tr}(\mathbf{S}_\tau^{-1} (\mathbf{Y} - \mathbf{Z}\mathbf{B} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})^\top \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{Z}\mathbf{B} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda}) \mathbf{S}_\tau^{-1} (\boldsymbol{\Omega}_d)^{-1})}. \end{aligned}$$

La posteriori condizionata non è riconducibile a nessuna densità nota per $\boldsymbol{\Omega}_d$, ma verrà utilizzata all'interno di uno *step Metropolis Hasting* all'interno dell'algoritmo di stima. Le distribuzioni a posteriori per $\boldsymbol{\beta}$ e per \mathbf{v} , ricavando gli elementi \mathbf{D} e $\boldsymbol{\Omega}$ da $\boldsymbol{\Omega}_d$, rimangono identiche a quelle viste precedentemente.

Per β si ha $\beta|\Omega_d, \mathbf{v} \sim \mathbf{N}(\hat{\boldsymbol{\mu}}_\beta^p, \hat{\boldsymbol{\Sigma}}_\beta^p)$, con:

$$\begin{aligned}\hat{\boldsymbol{\mu}}_\beta^p &= \text{vec}((\mathbf{Y} - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})^\top \mathbf{V}^{-1} \mathbf{Z}^\top (\mathbf{Z} \mathbf{V}^{-1} \mathbf{Z}^\top)^{-1}) \\ \hat{\boldsymbol{\Sigma}}_\beta^p &= ((\mathbf{Z} \mathbf{V}^{-1} \mathbf{Z}^\top)^{-1}) \otimes \mathbf{D} \boldsymbol{\Omega} \mathbf{D}.\end{aligned}$$

mentre per \mathbf{v} si ha come distribuzione a posteriori sui singoli elementi $v_i|\beta, \Omega_d \sim \text{Gig}(\frac{2-d}{2}, 2 + \boldsymbol{\lambda}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{\lambda}, \boldsymbol{\varepsilon}_i^\top (\mathbf{D} \boldsymbol{\Omega} \mathbf{D})^{-1} \boldsymbol{\varepsilon}_i)$, $\boldsymbol{\varepsilon}_i = \mathbf{y}_i - \mathbf{Z}_i \mathbf{B}$.

3.3 VAR quantilico con restrizioni

In questa sezione viene mostrato come estendere al caso quantilico la stima di un modello VAR con restrizioni, come presentato in 2.2. Assumiamo un set di restrizioni lineari su β del tipo:

$$\beta = \text{vec}(\mathbf{B}) = \mathbf{C} \boldsymbol{\delta} + \mathbf{c}.$$

Come spiegato nella sezione 2.2, questa forma è in grado di offrire sia una restrizione sul numero di parametri, sia di introdurre restrizioni sulla forma di β e sulle relazioni tra i suoi elementi. Sostituendo a β la forma in $\boldsymbol{\delta}$, all'interno del modello per $\mathbf{y} = \text{vec}(\mathbf{Y})$ visto in (3.3), si ha:

$$\mathbf{y} = \text{vec}(\mathbf{Y}) = \mathbf{X}(\mathbf{C} \boldsymbol{\delta} + \mathbf{c}) + \mathbf{v} \otimes (\mathbf{D} \boldsymbol{\lambda}) + (\sqrt{\mathbf{v}} \otimes (\mathbf{D} \boldsymbol{\Omega}^{\frac{1}{2}})) \mathbf{u}, \quad (3.12)$$

riscrivibile per comodità, dato che \mathbf{X} e \mathbf{c} sono note, come:

$$\mathbf{y}^* = \mathbf{y} - \mathbf{X} \mathbf{c} = (\mathbf{X} \mathbf{C}) \boldsymbol{\delta} + \mathbf{v} \otimes (\mathbf{D} \boldsymbol{\lambda}) + (\sqrt{\mathbf{v}} \otimes (\mathbf{D} \boldsymbol{\Omega}^{\frac{1}{2}})) \mathbf{u}.$$

Definiamo l'operatore *invvec*, tale che $\text{invvec}(\text{vec}(A)) = A$, e definiamo $\mathbf{Y}^* = \text{invvec}(\mathbf{y} - \mathbf{X} \mathbf{c})$, \mathbf{Y}^* matrice $n \times d$. Per la stima bayesiana risulta necessario specificare una priori per $\boldsymbol{\delta}$; ho scelto una priori normale, come introdotta precedentemente per β . Di seguito riporto tutte le priori scelte per $(\boldsymbol{\delta}, \boldsymbol{\Omega}, \mathbf{D}, \mathbf{v})$:

$$\begin{aligned}\pi(v_i) &= e^{-v_i} \\ \pi(\boldsymbol{\delta}) &= |2\pi \boldsymbol{\Sigma}_\delta|^{-\frac{1}{2}} e^{-(\boldsymbol{\delta} - \boldsymbol{\mu}_\delta)^\top (2\boldsymbol{\Sigma}_\delta)^{-1} (\boldsymbol{\delta} - \boldsymbol{\mu}_\delta)} \\ \pi(\boldsymbol{\Omega}) &\propto |\mathbf{C} \boldsymbol{\Omega}|^{-(c-d-1)} e^{-\frac{\text{Tr}(\mathbf{C} \boldsymbol{\Omega}^{-1})}{2}}\end{aligned}$$

$$\pi(\mathbf{D}) \propto 1.$$

Date le priori, la distribuzione congiunta del modello VAR(p) con restrizioni risulta la seguente:

$$\begin{aligned} \pi(\mathbf{Y}, \boldsymbol{\delta}, \mathbf{D}, \boldsymbol{\Omega}, \mathbf{v} | \mathbf{X}) &\propto |\boldsymbol{\Sigma}|^{-1/2} \\ &\times e^{-\frac{1}{2}(\mathbf{y}^* - \mathbf{XC}\boldsymbol{\delta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})^\top (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{y}^* - \mathbf{XC}\boldsymbol{\delta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})} \\ &\times e^{-\sum_1^n v_i |\mathbf{C}\boldsymbol{\Omega}|^{-(c-d-1)}} e^{-\frac{\text{Tr}(\mathbf{C}\boldsymbol{\Omega}\boldsymbol{\Omega}^{-1})}{2}}. \end{aligned}$$

Dalla distribuzione congiunta del modello con restrizioni possiamo ricavare le distribuzioni condizionate sui singoli parametri di $(\boldsymbol{\delta}, \boldsymbol{\Omega}, \mathbf{D}, \mathbf{v})$. Di seguito viene derivata la distribuzione a posteriori condizionata per il parametro $\boldsymbol{\delta}$:

$$\begin{aligned} \pi(\boldsymbol{\delta} | \mathbf{D}, \boldsymbol{\Omega}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\frac{1}{2}(\mathbf{y}^* - \mathbf{XC}\boldsymbol{\delta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})^\top (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{y}^* - \mathbf{XC}\boldsymbol{\delta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})} \\ &\times e^{-\frac{1}{2}(\boldsymbol{\delta} - \boldsymbol{\mu}_\delta)^\top \boldsymbol{\Sigma}_\delta^{-1} (\boldsymbol{\delta} - \boldsymbol{\mu}_\delta)} \\ \pi(\boldsymbol{\delta} | \mathbf{D}, \boldsymbol{\Omega}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\frac{1}{2}\boldsymbol{\delta}^\top (\mathbf{XC})^\top (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} \mathbf{XC}\boldsymbol{\delta}} \\ &\times e^{\boldsymbol{\delta}^\top (\mathbf{XC})^\top (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{y}^* - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})} \\ &\times e^{-\frac{1}{2}\boldsymbol{\delta}^\top \boldsymbol{\Sigma}_\delta^{-1} \boldsymbol{\delta}} e^{\boldsymbol{\delta}^\top \boldsymbol{\Sigma}_\delta^{-1} \boldsymbol{\mu}_\delta} \\ \pi(\boldsymbol{\delta} | \mathbf{D}, \boldsymbol{\Omega}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\frac{1}{2}\boldsymbol{\delta}^\top ((\mathbf{XC})^\top (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} \mathbf{XC} + \boldsymbol{\Sigma}_\delta^{-1}) \boldsymbol{\delta}} \\ &\times e^{\boldsymbol{\delta}^\top ((\mathbf{XC})^\top (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{y}^* - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda}) + \boldsymbol{\Sigma}_\delta^{-1} \boldsymbol{\mu}_\delta)}, \end{aligned}$$

la posteriori di $\boldsymbol{\delta}$ risulta essere riconducibile ad una distribuzione nota, corrispondente a $\mathbf{N}(\hat{\boldsymbol{\mu}}_\delta^p, \hat{\boldsymbol{\Sigma}}_\delta^p)$:

$$\begin{aligned} \hat{\boldsymbol{\mu}}_\delta^p &= \hat{\boldsymbol{\Sigma}}_\delta^p ((\mathbf{XC})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}^* - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda}) + \boldsymbol{\Sigma}_\delta^{-1} \boldsymbol{\mu}_\delta) \\ \hat{\boldsymbol{\Sigma}}_\delta^p &= (\boldsymbol{\Sigma}_\delta^{-1} + (\mathbf{XC})^\top \boldsymbol{\Sigma}^{-1} \mathbf{XC})^{-1}. \end{aligned}$$

Ricordando la scrittura come prodotto di Kronecker e tramite l'operatore vettorizzazione, si può riscrivere il tutto come:

$$\begin{aligned} \hat{\boldsymbol{\mu}}_\delta^p &= \hat{\boldsymbol{\Sigma}}_\delta^p (\mathbf{C}^\top (\mathbf{Z}^\top \otimes \mathbf{I}_d) (\mathbf{V}^{-1} \otimes (\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1}) \text{vec}((\mathbf{Y}^* - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})^\top) + \boldsymbol{\Sigma}_\delta^{-1} \boldsymbol{\mu}_\delta) \\ \hat{\boldsymbol{\Sigma}}_\delta^p &= (\boldsymbol{\Sigma}_\delta^{-1} + \mathbf{C}^\top (\mathbf{Z} \otimes \mathbf{I}_d) (\mathbf{V}^{-1} \otimes (\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1}) (\mathbf{Z}^\top \otimes \mathbf{I}_d) \mathbf{C})^{-1}. \end{aligned}$$

Sfruttando le proprietà del prodotto di Kronecker e dell'operatore vettorizzazione si ottiene la seguente forma:

$$\begin{aligned} \hat{\boldsymbol{\mu}}_\delta^p &= \hat{\boldsymbol{\Sigma}}_\delta^p (\mathbf{C}^\top \text{vec}((\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{Y}^* - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})^\top \mathbf{V}^{-1} \mathbf{Z}^\top) + \boldsymbol{\Sigma}_\delta^{-1} \boldsymbol{\mu}_\delta) \\ \hat{\boldsymbol{\Sigma}}_\delta^p &= (\boldsymbol{\Sigma}_\delta^{-1} + \mathbf{C}^\top (\mathbf{Z}\mathbf{V}^{-1}\mathbf{Z}^\top) \otimes (\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} \mathbf{C})^{-1}. \end{aligned} \quad (3.13)$$

Per quanto riguarda \mathbf{v} , vettore dei fattori latenti, si ha che la distribuzione condizionata risulta proporzionale a:

$$\begin{aligned} \pi(\mathbf{V}|\boldsymbol{\delta}, \mathbf{D}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\sum_{i=1}^n v_i} |\mathbf{V}|^{-\frac{d}{2}} \\ &\times e^{-\frac{1}{2}(\mathbf{y}^* - \mathbf{XC}\boldsymbol{\delta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})^\top (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{y}^* - \mathbf{XC}\boldsymbol{\delta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})}. \end{aligned}$$

Sfruttando il fatto che \mathbf{V} è una matrice diagonale, e utilizzando *invvec* su $\mathbf{XC}\boldsymbol{\delta}$, è possibile riscrivere il modello come:

$$\begin{aligned} \pi(\mathbf{V}|\boldsymbol{\delta}, \mathbf{D}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\sum_{i=1}^n \frac{1}{2}(\mathbf{y}_i^* - \text{invvec}(\mathbf{XC}\boldsymbol{\delta})_i - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})^\top (v_i \boldsymbol{\Omega})^{-1} (\mathbf{y}_i^* - \text{invvec}(\mathbf{XC}\boldsymbol{\delta})_i - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})} \\ &\times \prod_{i=1}^n v_i^{-\frac{d}{2}} e^{-v_i} \\ \pi(\mathbf{V}|\boldsymbol{\delta}, \mathbf{D}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\sum_{i=1}^n \frac{1}{2}(\mathbf{y}_i^* - \text{invvec}(\mathbf{XC}\boldsymbol{\delta})_i)^\top (v_i \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\text{invvec}(\mathbf{XC}\boldsymbol{\delta})_i)} \\ &\times e^{-\sum_{i=1}^n \frac{1}{2}(v_i \mathbf{D}\boldsymbol{\lambda})^\top (v_i \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (v_i \mathbf{D}\boldsymbol{\lambda})} \\ &\times e^{-\sum_{i=1}^n \frac{1}{2}(\mathbf{y}_i^* - \text{invvec}(\mathbf{XC}\boldsymbol{\delta})_i)^\top (v_i \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{y}_i^* - \text{invvec}(\mathbf{XC}\boldsymbol{\delta})_i - \boldsymbol{\lambda} v_i \sqrt{\mathbf{D}}^\top)} \\ &\times \prod_{i=1}^n v_i^{-\frac{d}{2}} e^{-v_i} \\ \pi(\mathbf{V}|\boldsymbol{\delta}, \mathbf{D}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\sum_{i=1}^n \frac{1}{2v_i} (\mathbf{y}_i^* - \text{invvec}(\mathbf{XC}\boldsymbol{\delta})_i)^\top (\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{y}_i^* - \text{invvec}(\mathbf{XC}\boldsymbol{\delta})_i)} \\ &\times e^{-\sum_{i=1}^n \frac{v_i}{2} (\mathbf{D}\boldsymbol{\lambda})^\top (\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{D}\boldsymbol{\lambda})} \\ &\times \prod_{i=1}^n v_i^{\frac{(2-d)}{2} + 1} e^{-\frac{2v_i}{2}} \\ \pi(\mathbf{V}|\boldsymbol{\beta}, \mathbf{D}, \mathbf{v}, \mathbf{Y}, \mathbf{X}) &\propto \prod_{i=1}^n v_i^{\frac{(2-d)}{2} + 1} e^{-\frac{1}{2v_i} (\mathbf{y}_i^* - \text{invvec}(\mathbf{XC}\boldsymbol{\delta})_i)^\top (\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{y}_i^* - \text{invvec}(\mathbf{XC}\boldsymbol{\delta})_i)} \\ &\times e^{-\frac{v_i}{2} (2 + \boldsymbol{\lambda}^\top \mathbf{D} (\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{D}\boldsymbol{\lambda}))}, \end{aligned}$$

il singolo elemento v_i $i = 1, \dots, n$ risulta essere una Gaussiana inversa generalizzata, $v_i | \boldsymbol{\Omega}, \boldsymbol{\delta}, \mathbf{D}, \mathbf{Y}, \mathbf{X} \sim \text{Gig}(\frac{2-d}{2}, 2 + \boldsymbol{\lambda}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{\lambda}, \boldsymbol{\epsilon}_i^{*\top} (\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} \boldsymbol{\epsilon}_i^*)$, con $\boldsymbol{\epsilon}_i^* = \mathbf{Y}_i^* - \text{invvec}(\mathbf{XC}\boldsymbol{\delta})_i$ vettore dei residui di regressione all'istante i .

Per quanto riguarda la distribuzione a posteriori di $\boldsymbol{\Omega}$ si ha che la posteriori risulta proporzionale a:

$$\begin{aligned} \pi(\boldsymbol{\Omega}|\mathbf{v}, \boldsymbol{\delta}, \mathbf{D}, \mathbf{Y}, \mathbf{X}) &\propto |\boldsymbol{\Omega}|^{-t/2} \\ &\times e^{-\frac{1}{2}(\mathbf{y}^* - \mathbf{XC}\boldsymbol{\delta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})^\top \mathbf{V} \otimes \boldsymbol{\Omega}^{-1} (\mathbf{y}^* - \mathbf{XC}\boldsymbol{\delta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})} \\ &\times |\boldsymbol{\Omega}|^{-\frac{c+p+1}{2}} e^{-\frac{\text{Tr}(\mathbf{C}\boldsymbol{\Omega}^{-1})}{2}}. \end{aligned}$$

Per poter utilizzare l'operatore traccia sulla prima parte dell'equazione, così da ottenere una posteriori nota, risulta necessario utilizzare l'operatore *invec* su $\mathbf{XC}\boldsymbol{\delta}$:

$$\begin{aligned} \pi(\boldsymbol{\Omega}|\boldsymbol{v}, \boldsymbol{\beta}, \mathbf{D}, \mathbf{Y}, \mathbf{X}) &\propto |\boldsymbol{\Omega}|^{-t/2} \\ &\times e^{-\frac{1}{2}\text{tr}(\mathbf{D}^{-1}(\mathbf{Y}^* - \text{invec}(\mathbf{XC}\boldsymbol{\delta}) - \boldsymbol{v} \otimes \boldsymbol{\lambda}^T \mathbf{D})^T \mathbf{V}^{-1}(\mathbf{Y}^* - \text{invec}(\mathbf{XC}\boldsymbol{\delta}) - \boldsymbol{v} \otimes \boldsymbol{\lambda}^T \mathbf{D})\mathbf{D}^{-1}\boldsymbol{\Omega}^{-1})} \\ &\times |\boldsymbol{\Omega}|^{-\frac{c+p+1}{2}} e^{-\frac{\text{tr}(\mathbf{C}\boldsymbol{\Omega}\boldsymbol{\Omega}^{-1})}{2}}. \end{aligned}$$

Raccogliendo in $\boldsymbol{\Omega}$ si ha:

$$\begin{aligned} \pi(\boldsymbol{\Omega}|\boldsymbol{v}, \boldsymbol{\beta}, \mathbf{D}, \mathbf{Y}, \mathbf{X}) &\propto |\boldsymbol{\Omega}|^{-\frac{c+p+1+n}{2}} \\ &e^{-\frac{\text{tr}((\mathbf{C}\boldsymbol{\Omega} + \mathbf{D}^{-1}(\mathbf{Y}^* - \text{invec}(\mathbf{XC}\boldsymbol{\delta}) - \boldsymbol{v} \otimes \boldsymbol{\lambda}^T \mathbf{D})^T \mathbf{V}^{-1}(\mathbf{Y}^* - \text{invec}(\mathbf{XC}\boldsymbol{\delta}) - \boldsymbol{v} \otimes \boldsymbol{\lambda}^T \mathbf{D})\mathbf{D}^{-1})\boldsymbol{\Omega}^{-1})}{2}}. \end{aligned}$$

La distribuzione condizionata risulta quindi $\boldsymbol{\Omega}|\boldsymbol{v}, \boldsymbol{\delta}, \mathbf{D}, \mathbf{Y}, \mathbf{X} \sim \text{IW}(\widehat{\mathbf{C}}_{\boldsymbol{\Omega}}^p, \widehat{c}^p)$, con:

$$\begin{aligned} \widehat{\mathbf{C}}_{\boldsymbol{\Omega}}^p &= \mathbf{C}_{\boldsymbol{\Omega}} + \mathbf{D}^{-1}(\mathbf{Y}^* - \text{invec}(\mathbf{XC}\boldsymbol{\delta}) - \boldsymbol{v} \otimes \boldsymbol{\lambda}^T \mathbf{D})^T \mathbf{V}^{-1}(\mathbf{Y}^* - \text{invec}(\mathbf{XC}\boldsymbol{\delta}) - \boldsymbol{v} \otimes \boldsymbol{\lambda}^T \mathbf{D})\mathbf{D}^{-1} \\ \widehat{c}^p &= c + n. \end{aligned} \tag{3.14}$$

Per quanto riguarda \mathbf{D} , anche in questo caso non risulta possibile ricondursi a nessuna distribuzione nota; la distribuzione condizionata risulta essere:

$$\begin{aligned} \pi(\mathbf{D}|\boldsymbol{v}, \boldsymbol{\delta}, \boldsymbol{\Omega}, \mathbf{Y}, \mathbf{X}) &\propto |\mathbf{D}|^{-n} \\ &\times e^{-\frac{1}{2}(\mathbf{y}^* - \mathbf{XC}\boldsymbol{\delta} - \boldsymbol{v} \otimes \mathbf{D}\boldsymbol{\lambda})^T (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{y}^* - \mathbf{XC}\boldsymbol{\delta} - \boldsymbol{v} \otimes \mathbf{D}\boldsymbol{\lambda})}. \end{aligned}$$

3.4 Estensione *spike-and-slab*

Una possibile estensione del modello VAR quantilico precedentemente introdotto, che ci permetta di selezionare quegli elementi di \mathbf{B} che hanno un effetto maggiore sulla risposta, prevede l'utilizzo di priori di tipo *spike-and-slab* su \mathbf{B} .

Rispetto al modello presentato in 3.2, si è scelto di dividere $\boldsymbol{\beta}$ in $\boldsymbol{\mu}$, intercette del modello, e $\boldsymbol{\beta}$, effetto di covariate e ritardi di \mathbf{Y} , così da applicare le restrizioni delle priori *spike-and-slab* solamente sulla nuova $\boldsymbol{\beta}$, lasciando così il *drift* libero da restrizioni.

Il modello viene definito come:

$$\pi(\mathbf{y}_i | \mathbf{Z}_i, \boldsymbol{\mu}, \mathbf{D}, \mathbf{B}, v_i, \pi_0, \boldsymbol{\Gamma}, \mathbf{k}_1, \sigma_\varepsilon, \boldsymbol{\Omega}) = \phi_d(\boldsymbol{\mu} + \mathbf{Z}_i \mathbf{B} + \boldsymbol{\lambda} \sqrt{\mathbf{D}} v_i, \boldsymbol{\varsigma}^2 v_i \mathbf{D} \boldsymbol{\Sigma}_\tau \mathbf{R} \boldsymbol{\Sigma}_\tau \mathbf{D})$$

con $\mathbf{k}_1 = (k_{11}, k_{21}, \dots, k_{d1})$, $\sigma_\varepsilon = (\sigma_{\varepsilon_1}, \sigma_{\varepsilon_2}, \dots, \sigma_{\varepsilon_M})$, e con γ , con la stessa forma di \mathbf{B} , definita come:

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1(M-1)} & \gamma_{1M} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2(M-1)} & \gamma_{2M} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma_{d1} & \gamma_{d2} & \dots & \gamma_{d(M-1)} & \gamma_{dM} \end{bmatrix}.$$

Di seguito vengono riportate le densità a priori che verranno utilizzate:

$$\begin{aligned} \pi(\boldsymbol{\mu}) &\propto 1 \\ \pi(v_i) &= e^{-v_i} \\ \pi(\sigma_{\varepsilon_j}) &\propto \sigma_{\varepsilon_j}^{-a_\varepsilon-1} e^{-\frac{b_\varepsilon}{\sigma_{\varepsilon_j}}} \\ \pi(\mathbf{B}_{lj} | \sigma_{\varepsilon_j}, \gamma_{lj}, k_{l1}) &= N(0, \sigma_{\mathbf{B}_{lj}}) \\ \sigma_{\mathbf{B}_{lj}} &= k_{l1} \sigma_{\varepsilon_j} ((1 - \gamma_{lj}) k_0 + \gamma_{lj}), \\ 0 \leq k_0 &\ll \min(k_{l1}), \quad j = 1, 2, \dots, m \quad l = 1, 2, \dots, d \\ \gamma_{lj} &\sim \pi_0^{\gamma_i} (1 - \pi_0)^{1-\gamma_i} \\ \pi(\pi_{0l}) &\propto \pi_{0l}^{a_\pi-1} (1 - \pi_{0l})^{b_\pi} \\ \pi(k_{1i}) &= \propto k_{l1}^{-a_k-1} e^{-\frac{b_k}{k_{l1}}} \\ \pi(\mathbf{D}) &\propto 1 \\ \boldsymbol{\Omega} &\sim \text{IW}(c, \mathbf{C}). \end{aligned}$$

La costruzione di $\boldsymbol{\beta} = \text{vec}(\mathbf{B})$ avviene in maniera gerarchica; il primo elemento ad essere generato è π_{0l} , elemento compreso tra 0 ed 1, che regola a priori la probabilità, per gli elementi appartenenti alla l -esima riga di \mathbf{B} di essere generati dalla *spike* o dalla componente *slab*. Il parametro γ_{lj} regola la generazione del parametro \mathbf{B}_{lj} , regolandone l'appartenenza alla componente *spike* (bassa variabilità $k_0 k_{l1} \sigma_{\varepsilon_j}$, con k_0 valore molto piccolo) o alla componente *slab* (alta variabilità $k_{l1} \sigma_{\varepsilon_j}$). Il parametro k_{l1} rappresenta la variabilità della componente *slab*, e viene generato diverso per ogni riga di \mathbf{B} , così da adattarsi alla scala della risposta. k_{l1} viene generato da una distribuzione gamma inversa, per ogni l , così da fornire tenere conto della scala relativa

alla l -esima variabile risposta nella costruzione dei $\mathbf{B}_{l\cdot}$. La componente σ_{ε_j} viene inserita all'interno della varianza di \mathbf{B}_{lj} così da tenere conto anche della scala della covariata j nella determinazione della varianza di \mathbf{B}_{lj} .

La distribuzione congiunta del modello risulta essere:

$$\begin{aligned}
\pi(\mathbf{y}, \mathbf{D}, \boldsymbol{\beta}, \mathbf{v}, \Gamma, \mathbf{k}_1, \boldsymbol{\sigma}_\varepsilon, \boldsymbol{\Omega} | \mathbf{X}) &\propto |\boldsymbol{\zeta} \mathbf{V} \otimes \mathbf{D} \mathbf{S}_\tau \boldsymbol{\Omega} \mathbf{S}_\tau \mathbf{D}|^{-1/2} \\
&\times \mathbf{e}^{-\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})^\top (\mathbf{V} \otimes \mathbf{D} \mathbf{S}_\tau \boldsymbol{\Omega} \mathbf{S}_\tau \mathbf{D})^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})} \\
&\times \mathbf{e}^{-\sum_{i=1}^n v_i} \mathbf{e}^{-\frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)^\top \boldsymbol{\Sigma}_\beta^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)} \\
&\times \prod_{l=1}^d \prod_{j=1}^M \pi_{0l}^{\gamma_{lj}} (1 - \pi_{0l})^{1 - \gamma_{lj}} \prod_{l=1}^d k_{l1}^{-a_k - 1} \mathbf{e}^{-\frac{b_k}{k_{l1}}} \\
&\times \prod_{l=1}^d \pi_{0l}^{a_\pi - 1} (1 - \pi_{0l})^{b_\pi - 1} \prod_{j=1}^M \mathbf{e}^{-\frac{b_\varepsilon}{\sigma_{\varepsilon_j}}} \sigma_{\varepsilon_j}^{-a_\varepsilon - 1} \\
&\times |\mathbf{C}_\boldsymbol{\Omega}|^{-(c-d-1)} \mathbf{e}^{-\frac{\text{Tr}(\mathbf{C}_\boldsymbol{\Omega} \boldsymbol{\Omega}^{-1})}{2}}.
\end{aligned} \tag{3.15}$$

Dalla densità congiunta appena scritta possiamo determinare le distribuzioni a posteriori condizionate per i vari parametri di interesse, che verranno successivamente utilizzate per la stima tramite algoritmo MCMC. La distribuzione a posteriori di \mathbf{v} non subisce variazioni rispetto a quella presentata in 2.1; $v_i \sim \text{GIG}(\frac{2-d}{2}, 2 + \boldsymbol{\lambda}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{\lambda}, \boldsymbol{\varepsilon}_i^\top (\mathbf{D} \mathbf{S}_\tau \mathbf{R} \mathbf{S}_\tau \mathbf{D})^{-1} \boldsymbol{\varepsilon}_i)$, con $\boldsymbol{\varepsilon}_i = \mathbf{y}_i - \boldsymbol{\mu} - \mathbf{Z}_i \mathbf{B}$. La distribuzione a posteriori di \mathbf{D} non risulta riconducibile a nessuna forma nota. La distribuzione per $\boldsymbol{\mu}$ risulta proporzionale a:

$$\begin{aligned}
\pi(\boldsymbol{\mu} | \dots) &\propto \mathbf{e}^{-\frac{1}{2}(\mathbf{y} - \mathbf{1}_n \otimes \boldsymbol{\mu} - \mathbf{X}\boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})^\top (\mathbf{V} \otimes \mathbf{S}_\tau \boldsymbol{\Omega} \mathbf{S}_\tau \mathbf{D})^{-1} (\mathbf{y} - \mathbf{1}_n \otimes \boldsymbol{\mu} - \mathbf{X}\boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})} \\
&\propto \mathbf{e}^{-\frac{1}{2}(\mathbf{1}_n \otimes \boldsymbol{\mu})^\top (\mathbf{V} \otimes \mathbf{D} \mathbf{S}_\tau \boldsymbol{\Omega} \mathbf{S}_\tau \mathbf{D})^{-1} (\mathbf{1}_n \otimes \boldsymbol{\mu})} \mathbf{e}^{-\frac{1}{2}(\mathbf{1}_n \otimes \boldsymbol{\mu})^\top (\mathbf{V} \otimes \mathbf{D} \mathbf{S}_\tau \boldsymbol{\Omega} \mathbf{S}_\tau \mathbf{D})^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})}.
\end{aligned}$$

La posteriori risulta quindi $\boldsymbol{\mu} | \dots \sim \mathbf{N}(\boldsymbol{\mu}^p, \boldsymbol{\Sigma}^p)$, con:

$$\begin{aligned}
\boldsymbol{\mu}^p &= \boldsymbol{\Sigma}^p (\mathbf{1}_n^\top \mathbf{V}^{-1} \otimes (\mathbf{D} \mathbf{S}_\tau \boldsymbol{\Omega} \mathbf{S}_\tau \mathbf{D})^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})) \\
\boldsymbol{\Sigma}^p &= (\mathbf{1}_n^\top \mathbf{V}^{-1} \mathbf{1}_n \otimes (\mathbf{D} \mathbf{S}_\tau \boldsymbol{\Omega} \mathbf{S}_\tau \mathbf{D})^{-1})^{-1}.
\end{aligned}$$

Semplificando si ottiene:

$$\begin{aligned}
\boldsymbol{\mu}^p &= \frac{(\mathbf{v}^{-1} (\mathbf{Y} - \mathbf{Z}\mathbf{B} - \mathbf{v} \mathbf{D}\boldsymbol{\lambda}))}{\sum_{i=1}^n \frac{1}{v_i}} \\
\boldsymbol{\Sigma}^p &= \left(\sum_{i=1}^n \frac{1}{v_i} \right) \mathbf{D} \mathbf{S}_\tau \boldsymbol{\Omega} \mathbf{S}_\tau \mathbf{D}.
\end{aligned}$$

La distribuzione a posteriori di Ω risulta $\Omega | \dots \sim \text{IW}(\widehat{\mathbf{C}}_\Omega^p, \widehat{c}^p)$, con:

$$\begin{aligned}\widehat{\mathbf{C}}_\Omega^p &= \mathbf{C}_\Omega + \mathbf{D}^{-1}(\mathbf{Y} - \mathbf{1}_n \boldsymbol{\mu} - \mathbf{Z}^\top \mathbf{B} - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})^\top \mathbf{V}^{-1}(\mathbf{Y} - \mathbf{1}_n \boldsymbol{\mu} - \mathbf{Z}^\top \mathbf{B} - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D}) \mathbf{D}^{-1} \\ \widehat{c}^p &= c + n.\end{aligned}$$

La distribuzione a posteriori di β non subisce variazioni rispetto a quella vista in 3.2, e risulta essere $\beta | \dots \sim \text{N}(\widehat{\boldsymbol{\mu}}_\beta^p, \widehat{\boldsymbol{\Sigma}}_\beta^p)$, con:

$$\begin{aligned}\widehat{\boldsymbol{\mu}}_\beta^p &= \widehat{\boldsymbol{\Sigma}}_\beta^p (\text{vec}(\Omega^{-1}(\mathbf{Y} - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})^\top \mathbf{V}^{-1} \mathbf{Z}^\top) + \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\mu}_\beta) \\ \widehat{\boldsymbol{\Sigma}}_\beta^p &= (\boldsymbol{\Sigma}_\beta^{-1} + (\mathbf{Z} \mathbf{V}^{-1} \mathbf{Z}^\top) \otimes \Omega^{-1})^{-1}.\end{aligned}$$

Da notare il fatto che $\boldsymbol{\Sigma}_\beta$ non è più un iperparametro fissato ma sarà generato con elementi sulla diagonale *spike slab*, che influenzeranno quindi la posteriori.

La distribuzione a posteriori per k_{l1} , variabile di scala relativa alla l -esima riga di \mathbf{B} , risulta:

$$\begin{aligned}\pi(\mathbf{k}_1 | \dots) &\propto \prod_{l=1}^d k_{l1}^{-a_k-1} \mathbf{e}^{-\frac{b_k}{k_{l1}}} \\ &\quad \times |\boldsymbol{\Sigma}_\beta|^{-\frac{1}{2}} \mathbf{e}^{-\frac{1}{2}(\beta - \boldsymbol{\mu}_\beta)^\top \boldsymbol{\Sigma}_\beta^{-1}(\beta - \boldsymbol{\mu}_\beta)} \\ &\propto \prod_{l=1}^d k_{l1}^{-a_k-1} \mathbf{e}^{-\frac{b_k}{k_{l1}}} \prod_{j=1}^M |k_{l1} \sigma_{\varepsilon_j} (k_0(1 - \gamma_{lj}) + \gamma_{lj})|^{-\frac{1}{2}} \mathbf{e}^{-\frac{\mathbf{B}_{lk}^2}{2\sigma_{\varepsilon_j} k_{l1} (k_0(1 - \gamma_{lj}) + \gamma_{lj})}} \\ &\propto \prod_{l=1}^d k_{l1}^{-a_k-1} \mathbf{e}^{-\frac{b_k}{k_{l1}}} |(k_{l1})|^{-\frac{M}{2}} \mathbf{e}^{-\frac{1}{2k_{l1}} \sum_{j=1}^M \frac{\mathbf{B}_{lk}^2}{\sigma_{\varepsilon_j} ((1 - \gamma_{lk})k_0 + \gamma_{lk})}} \\ &\propto \prod_{l=1}^d k_{l1}^{-a_k-1 - \frac{M}{2}} \mathbf{e}^{-\frac{b_k}{k_{l1}}} \mathbf{e}^{-\frac{1}{2k_{l1}} \sum_{j=1}^M \frac{\mathbf{B}_{lk}^2}{\sigma_{\varepsilon_j} ((1 - \gamma_{lk})k_0 + \gamma_{lk})}}.\end{aligned}$$

Di conseguenza si ha che la distribuzione a posteriori condizionata coincide quindi con una densità nota: $k_{l1} | \dots \sim \text{IG}(a_k + \frac{M}{2}, b_k + \frac{1}{2} \sum_{j=1}^M \frac{\mathbf{B}_{lk}^2}{\sigma_{\varepsilon_j} ((1 - \gamma_{lk})k_0 + \gamma_{lk})})$ per $l = 1, \dots, d$. La posteriori di γ_{lj} , variabile indicatrice che regola l'assegnazione dell' elemento (l, j) di \mathbf{B} alla componente *slab* o alla componente

spike, risulta proporzionale a:

$$\begin{aligned}
\pi(\mathbf{\Gamma} | \dots) &\propto \prod_{l=1}^d \prod_{j=1}^M \pi_{0l}^{\gamma_{lj}} (1 - \pi_{0l})^{1-\gamma_{lj}} \\
&\quad \times |\boldsymbol{\Sigma}_{\beta}|^{-\frac{1}{2}} \mathbf{e}^{-\frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\mu}_{\beta})^{\top} \boldsymbol{\Sigma}_{\beta}^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}_{\beta})} \\
&\propto \prod_{l=1}^d \prod_{j=1}^M \pi_{0l}^{\gamma_{lj}} (1 - \pi_{0l})^{1-\gamma_{lj}} \\
&\quad \times |k_{l1} \sigma_{\varepsilon_j} (k_0(1 - \gamma_{lj}) + \gamma_{lj})|^{-\frac{1}{2}} \mathbf{e}^{-\frac{\mathbf{B}_{lj}^2}{2k_{l1} \sigma_{\varepsilon_j} (k_0(1 - \gamma_{lj}) + \gamma_{lj})}} \\
&\propto \prod_{l=1}^d \prod_{j=M}^k \frac{\pi_{0l}^{\gamma_{lj}} (1 - \pi_{0l})^{1-\gamma_{lj}}}{\sqrt{k_{l1} \sigma_{\varepsilon_j} (k_0(1 - \gamma_{lj}) + \gamma_{lj})}} \mathbf{e}^{-\frac{\mathbf{B}_{lj}^2}{2k_{l1} \sigma_{\varepsilon_j} (k_0(1 - \gamma_{lj}) + \gamma_{lj})}}.
\end{aligned}$$

La distribuzione risulta, come la priori, una binomiale, con nuove probabilità di appartenere alla componente *slab* o alla componente *spike* proporzionali a:

$$\begin{aligned}
\pi(\gamma_{lj} = 1 | \dots) &\propto \frac{\pi_{0l}}{\sqrt{k_{l1} \sigma_{\varepsilon_j}}} \mathbf{e}^{-\frac{\mathbf{B}_{lj}^2}{2k_{l1} \sigma_{\varepsilon_j}}} \\
\pi(\gamma_{lj} = 0 | \dots) &\propto \frac{1 - \pi_{0l}}{\sqrt{k_0 k_{l1} \sigma_{\varepsilon_j}}} \mathbf{e}^{-\frac{\mathbf{B}_{lj}^2}{2k_0 k_{l1} \sigma_{\varepsilon_j}}}.
\end{aligned}$$

La distribuzione a posteriori per σ_{ε_j} , variabile di scala per le colonna j di \mathbf{B} risulta essere:

$$\begin{aligned}
\pi(\sigma_{\varepsilon_j} | \dots) &\propto \mathbf{e}^{-\frac{\boldsymbol{\beta}^{\top} \boldsymbol{\Sigma}_{\beta}^{-1} \boldsymbol{\beta}}{2\sigma_{\varepsilon_j}}} \sigma_{\varepsilon_j}^{-(a_{\varepsilon}+1)} \mathbf{e}^{-\frac{b_{\varepsilon}}{\sigma_{\varepsilon_j}}} \\
&\propto \sigma_{\varepsilon_j}^{-(a_{\varepsilon}+1)} \mathbf{e}^{-\frac{b_{\varepsilon}}{\sigma_{\varepsilon_j}}} \\
&\quad \times \prod_{l=1}^d |2k_{l1} \sigma_{\varepsilon_j} (k_0(1 - \gamma_{lj}) + \gamma_{lj})|^{-\frac{1}{2}} \mathbf{e}^{-\frac{\mathbf{B}_{lj}^2}{2k_{l1} \sigma_{\varepsilon_j} (k_0(1 - \gamma_{lj}) + \gamma_{lj})}} \\
&\propto \sigma_{\varepsilon_j}^{-(a_{\varepsilon}+1)} \mathbf{e}^{-\frac{b_{\varepsilon}}{\sigma_{\varepsilon_j}}} \\
&\quad \times |\sigma_{\varepsilon_j}|^{-\frac{d}{2}} \mathbf{e}^{-\sum_{l=1}^d \frac{\mathbf{B}_{lj}^2}{2k_{l1} \sigma_{\varepsilon_j} (k_0(1 - \gamma_{lj}) + \gamma_{lj})}} \\
&\propto \mathbf{e}^{-\frac{1}{\sigma_{\varepsilon_j}} (\sum_{l=1}^d \frac{\mathbf{B}_{lj}^2}{2k_{l1} (k_0(1 - \gamma_{lj}) + \gamma_{lj})} + b_{\varepsilon})} \sigma_{\varepsilon_j}^{-a_{\varepsilon} - 1 - \frac{d}{2}},
\end{aligned}$$

che corrisponde ad una distribuzione $\sigma_{\varepsilon_j} | \dots \sim \text{IG}(a_{\varepsilon} + \frac{d}{2}, \sum_{l=1}^d \frac{\mathbf{B}_{lj}^2}{2k_{l1} (k_0(1 - \gamma_{lj}) + \gamma_{lj})} + b_{\varepsilon})$ per $j = 1, \dots, M$.

La distribuzione a posteriori per π_{0l} , probabilità di appartenere alla componente *slab* per gli elementi della l -esima riga di \mathbf{B} , risulta essere:

$$\begin{aligned}\pi(\pi_{0l}|\dots) &\propto \pi_{0l}^{a_\pi-1} (1-\pi_{0l})^{b_\pi-1} \pi_{0l}^{\sum_{j=1}^M \gamma_{lj}} (1-\pi_{0l})^{\sum_{j=1}^M (1-\gamma_{lj})} \\ &\propto \pi_{0l}^{a_\pi-1+\sum_{j=1}^M \gamma_{lj}} (1-\pi_{0l})^{b_\pi-1+\sum_{j=1}^M (1-\gamma_{lj})},\end{aligned}$$

che corrisponde ad una posteriori condizionata nota $\pi_{0l}|\dots \sim \text{Beta}(a_\pi + \sum_{j=1}^M \gamma_{lj}, b_\pi + \sum_{j=1}^M (1-\gamma_{lj}))$, per $l = 1, \dots, d$.

3.5 Analisi Strutturali quantiliche

In questa sezione presento la mia proposta per estendere al modello VAR quantilico bayesiano la stima di *impulse response function* e *forecast error variance decomposition*. Un modello VAR quantilico frequentista soffre la mancanza della stima diretta della matrice di covarianza; per la stima della matrice di covarianza relativa ad un modello VAR quantilico stimato l'unico metodo proposto è quello di Cecchetti e L. Li 2008, che è uno strumento da usare successivamente alla stima del modello. Il vantaggio della stima tramite MCMC di un modello VAR quantilico bayesiano sta nella stima diretta, tramite la generazione di valori, della distribuzione della matrice di covarianza nel quantile τ di interesse. A differenza della stima puntuale della matrice di varianza ottenuta da Cecchetti e L. Li 2008, la stima del modello tramite MCMC offre la possibilità di stimare l'intera distribuzione della matrice di covarianza. La conoscenza della distribuzione della matrice di covarianza, oltre a quella dei coefficienti di regressione, può quindi permettermi non solo di avere una stima puntuale della matrice di covarianza, e di *impulse response function* e *forecast error variance decomposition*, ma di ricavare delle stime delle distribuzioni anche di queste quantità, utilizzabili per definire intervalli ed effettuare test. Di seguito viene mostrata la derivazione di *impulse response function* e *forecast error variance decomposition* quantiliche.

3.5.1 Impulse response function quantiliche

L'interpretazione dei risultati di un VAR risulta spesso complicata per via delle interazioni tra le variabili sia nei termini autoregressivi che nelle innovazioni. Le IRF e le GIRF possono aiutare nella comprensione e interpretazione dei risultati, fornendoci una idea dell'effetto nel tempo che una variabile ha sulle altre. Di seguito mostro come ho derivato, sfruttando la rappresentazione di un VAR quantilico come mistura di Normali, una stima delle IRF e GIRF quantiliche. La procedura di derivazione risulta simile a quanto mostrato in 2.3.1. Ricordando la forma del modello stimato VAR(p):

$$\mathbf{y}_i = \mathbf{B}\mathbf{Z}_i + v_i\mathbf{D}\boldsymbol{\lambda} + \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}}\boldsymbol{\varepsilon}_i,$$

con $\boldsymbol{\varepsilon}_i \sim \mathbf{N}_d(0, \mathbf{I}_d)$, scomponiamo \mathbf{B} e \mathbf{Z}_i per sottolineare le componenti autoregressive:

$$\mathbf{y}_i = \boldsymbol{\mu} + \mathbf{A}_1\mathbf{y}_{i-1} + \cdots + \mathbf{A}_p\mathbf{y}_{i-p} + \mathbf{B}_x\mathbf{X}_i + v_i\mathbf{D}\boldsymbol{\lambda} + \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}}\boldsymbol{\varepsilon}_i.$$

Riscrivo il modello in forma canonica:

$$\begin{aligned} \begin{bmatrix} \mathbf{y}_i \\ \mathbf{y}_{i-1} \\ \vdots \\ \mathbf{y}_{i-p+1} \end{bmatrix} &= \begin{bmatrix} \boldsymbol{\mu} + \mathbf{B}\mathbf{x}_i \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_{p-1} & \mathbf{A}_p \\ \mathbf{I}_d & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{I}_d & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y}_{i-1} \\ \mathbf{y}_{i-2} \\ \vdots \\ \mathbf{y}_{i-p} \end{bmatrix} + \\ &+ \begin{bmatrix} v_i\mathbf{D}\boldsymbol{\lambda} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \sqrt{v_i}\mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}}\boldsymbol{\varepsilon}_i \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \end{aligned}$$

Il modello in forma canonica, può essere considerato come un VAR(1) in \mathbf{y}_i^c

$$\mathbf{y}_i^c = \begin{bmatrix} \boldsymbol{\mu} + \mathbf{B}\mathbf{x}_i + v_i\mathbf{D}\boldsymbol{\lambda} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_{p-1} & \mathbf{A}_p \\ \mathbf{I}_d & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{I}_d & 0 \end{bmatrix} \mathbf{y}_{i-1}^c + \begin{bmatrix} \sqrt{v_i}\mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}}\boldsymbol{\varepsilon}_i \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\mathbf{y}_i^c = \boldsymbol{\mu}_i^c + \mathbf{A}\mathbf{y}_{i-1}^c + \boldsymbol{\varepsilon}_i^c.$$

Si può riscrivere il modello VAR(1) in forma canonica per \mathbf{y}_i^c nella sua forma *VMA* :

$$\mathbf{y}_i^c = \boldsymbol{\mu}_i^c + \sum_{j=1}^{i-1} \mathbf{A}^j \boldsymbol{\mu}_{i-j}^c + \mathbf{A}^i \mathbf{y}_0^c + \sum_{j=0}^{i-1} \mathbf{A}^j \boldsymbol{\varepsilon}_{i-j}^c.$$

Si ha che, dato $\mathbf{J} = [\mathbf{I}_d, 0, \dots, 0]$ vettore in $d \times dp$, premoltiplicando per la matrice in forma canonica, $\mathbf{J}\mathbf{y}_i^c = \mathbf{y}_i$, ci si riconduce al modello iniziale per \mathbf{y}_i :

$$\mathbf{y}_i = \mathbf{J}\boldsymbol{\mu}_i^c + \sum_{j=1}^{i-1} \mathbf{J}\mathbf{A}^j \boldsymbol{\mu}_{i-j}^c + \mathbf{J}\mathbf{A}^i \mathbf{y}_0 + \sum_{j=0}^{i-1} \mathbf{J}\mathbf{A}^j \mathbf{J}^\top \mathbf{J} \sqrt{v_i} \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}} \boldsymbol{\varepsilon}_{i-j}. \quad (3.16)$$

Assumendo $\mathbf{y}_0 = 0$, e scrivendo con $\boldsymbol{\mu}_i^* = \mathbf{J}(\boldsymbol{\mu}_i^c + \sum_{j=1}^{i-1} \mathbf{A}^j \boldsymbol{\mu}_{i-j}^c)$, e $\Phi_j = \mathbf{J}\mathbf{A}^j \mathbf{J}^\top$ il modello risulta:

$$\mathbf{y}_i = \boldsymbol{\mu}_i^* + \sum_{j=0}^{i-1} \Phi_j \sqrt{v_{i-j}} \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}} \boldsymbol{\varepsilon}_{i-j}.$$

Qualora sia disponibile un ordinamento possiamo quindi stimare le IRF quantiliche a ritardo i , con impulso unitario sulla variabile j come:

$$QIRF_{ij} = \Phi_i \sqrt{v_0} \mathbf{D}^{-1} \boldsymbol{\Omega}^{-\frac{1}{2}} \mathbf{e}_j = \mathbf{J}\mathbf{A}^i \mathbf{J}^\top \sqrt{v_0} \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}} \mathbf{e}_j,$$

con \mathbf{e}_j vettore di lunghezza d con elemento pari a 1 per il j -esimo elemento e 0 altrove.

Operando in un contesto in cui non vi è un vero e proprio ordinamento tra variabili, si possono ricavare le GIRF quantiliche come:

$$GQIRF_{ij} = E(\mathbf{y}_i - \boldsymbol{\mu}_i^* | \boldsymbol{\varepsilon}_0 = \sqrt{v_0} \mathbf{D}\boldsymbol{\Omega} \mathbf{D} \mathbf{e}_j \mathbf{D}_{jj}^{-1}) = \mathbf{J}\mathbf{A}^i \mathbf{J}^\top \sqrt{v_0} \mathbf{D}\boldsymbol{\Omega} \mathbf{D} \mathbf{e}_j (\zeta_j \mathbf{D}_{jj})^{-1}.$$

Si può notare come sia le IRF quantiliche che le GIRF quantiliche dipendono da un elemento $\sqrt{v_0}$ ignoto, corrispondente in teoria al valore del fattore latente nell'istante 0 in cui avviene l'impulso. Una soluzione per la stima di v_0 , risulta quella di sostituire a v_0 il suo valore atteso marginale, o, operativamente, la media marginale \boldsymbol{v}^m degli v_i stimati.

Anche in questo caso, come visto nel caso di un VAR classico in 2.3.1, si avrà che le IRF quantiliche e le GIRF quantiliche così definite coincideranno tra loro nel caso in cui $j = 1$, per impulsi sulla prima variabile del sistema.

Alternativamente si può considerare la scrittura del modello con innovazioni Laplace Asimmetriche, eliminando così completamente il fattore latente v_0 dalle IRF. Possiamo riscrivere il nostro modello VARX(p) come:

$$\mathbf{y}_i = \boldsymbol{\mu} + \mathbf{A}_1 \mathbf{y}_{i-1} + \cdots + \mathbf{A}_p \mathbf{y}_{i-p} + \mathbf{B}_x \mathbf{X}_i + \mathbf{u}_i$$

con $\mathbf{u}_i \sim \text{AL}_d(0, \mathbf{D}\boldsymbol{\Omega}\mathbf{D})$. Sfruttando le proprietà di posizione-scala della distribuzione Laplace Asimmetrica, si può riscrivere il modello come:

$$\mathbf{y}_i = \boldsymbol{\mu} + \mathbf{A}_1 \mathbf{y}_{i-1} + \cdots + \mathbf{A}_p \mathbf{y}_{i-p} + \mathbf{B}_x \mathbf{X}_i + \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}} \mathbf{u}_i^0,$$

con $\mathbf{u}_i \sim \text{AL}_d(0, \mathbf{I}_d)$. Possiamo riscriverlo in forma VMA come:

$$\mathbf{y}_i = \boldsymbol{\mu}_i^* + \sum_{j=0}^{i-1} \Phi_j \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}} \mathbf{u}_{i-j}^0,$$

con $\boldsymbol{\mu}_i^* = \mathbf{J}(\boldsymbol{\mu}_i^c + \sum_{j=1}^{i-1} \mathbf{A}^j \boldsymbol{\mu}_{i-j}^c)$, $\boldsymbol{\mu}_i^c = (\boldsymbol{\mu} + \mathbf{B}_x \mathbf{X}_i, 0, \dots, 0)^\top$, $\boldsymbol{\mu}_i^c \in \mathbb{R}^{dp}$. A questo punto risulta immediato ricavare le IRF, come fatto precedentemente, come:

$$QIRF_i = \Phi_i \mathbf{D}^{-1} \boldsymbol{\Omega}^{-\frac{1}{2}} \mathbf{e}_j = \mathbf{J} \mathbf{A}^i \mathbf{J}^\top \mathbf{D} \boldsymbol{\Omega}^{\frac{1}{2}} \mathbf{e}_j,$$

e le GQIRF come:

$$E(\mathbf{y}_i - \boldsymbol{\mu}_i^* | \varepsilon_0 = \mathbf{D}\boldsymbol{\Omega}\mathbf{D}\mathbf{e}_j \mathbf{D}_{jj}^{-1}) = \mathbf{J} \mathbf{A}^i \mathbf{J}^\top \mathbf{D} \boldsymbol{\Omega} \mathbf{D} \mathbf{e}_j (\varsigma_j \mathbf{D}_{jj})^{-1}.$$

3.5.2 FEVD quantilica

Analogamente a quanto fatto per le IRF, è possibile derivare una forma esplicita per la stima della *forecast error variance decomposition* relativa al modello stimato, sfruttando la distribuzione normale di $\mathbf{y}|\mathbf{v}$. Riprendendo la forma VMA per il nostro modello vista in (3.16):

$$\mathbf{y}_i = \mathbf{J} \boldsymbol{\mu}_i^c + \sum_{j=1}^{i-1} \mathbf{J} \mathbf{A}^j \boldsymbol{\mu}_{i-j}^c + \mathbf{J} \mathbf{A}^i \mathbf{y}_0 + \sum_{j=0}^{i-1} \mathbf{J} \mathbf{A}^j \mathbf{J}^\top \mathbf{J} \sqrt{v_i} \mathbf{D} \boldsymbol{\Omega}^{\frac{1}{2}} \varepsilon_{i-j},$$

possiamo definire la previsione h passi in avanti come:

$$\mathbf{y}_{i+h} = \mathbf{y}_i + \sum_{j=0}^{h-1} \mathbf{J} \mathbf{A}^j \boldsymbol{\mu}_{i+j+1}^c + \sum_{j=0}^{h-1} \mathbf{J} \mathbf{A}^j \mathbf{J}^\top \mathbf{J} \sqrt{v_i} \mathbf{D} \boldsymbol{\Omega}^{\frac{1}{2}} \varepsilon_{i+j+1},$$

in cui solo la componente $\sum_{j=0}^{h-1} \mathbf{J}\mathbf{A}^j \mathbf{J}^\top \mathbf{J} \sqrt{v_i} \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}} \varepsilon_{i+j+1}$ dipende dalle innovazioni. Il valore atteso $E(\mathbf{y}_{i+h}|\mathbf{y}_i)$ risulta:

$$E(\mathbf{y}_{i+h}|\mathbf{y}_i) = \hat{\boldsymbol{\mu}}_{i+h} = \mathbf{y}_i + \sum_{j=0}^{h-1} \mathbf{J}\mathbf{A}^j \boldsymbol{\mu}_{i+j+1}^c.$$

Concentrandoci sull'errore relativo al l -esimo elemento di \mathbf{y}_{i+h} si ha che:

$$\mathbf{y}_{l,i+h} - \hat{\boldsymbol{\mu}}_{i+h} = \sum_{j=0}^{h-1} \sum_{k=1}^d (\mathbf{e}_l^\top \mathbf{J}\mathbf{A}^j \mathbf{J}^\top \mathbf{J} \sqrt{v_i} \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}} \varepsilon_{k,i+j+1}).$$

La varianza delle previsioni sulla l -esima variabile risulta quindi:

$$E((\mathbf{y}_{l,i+h} - m\hat{u}_{i+h})^2) = \sum_{j=0}^{h-1} \sum_{k=1}^d (\mathbf{e}_l^\top \mathbf{J}\mathbf{A}^j \mathbf{J}^\top \mathbf{J} \sqrt{v_i} \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}} \mathbf{e}_k)^2.$$

A questo punto è possibile definire una *quantile forecast error variance decomposition* sulla variabile l , che misura l'input, h passi in avanti, alla varianza sul quantile τ_l dato dalle innovazioni sulla variabile k , come:

$$QFEVD_{lk,h} = \sum_{j=0}^{h-1} (\mathbf{e}_l^\top \mathbf{J}\mathbf{A}^j \mathbf{J}^\top \mathbf{J} \sqrt{v_i} \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}} \mathbf{e}_k)^2 / \sum_{j=0}^{h-1} \sum_{k=1}^d (\mathbf{e}_l^\top \mathbf{J}\mathbf{A}^j \mathbf{J}^\top \mathbf{J} \sqrt{v_i} \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}} \mathbf{e}_k)^2.$$

Analogamente è possibile definire delle *generalised quantile forecast error variance decomposition* come:

$$\begin{aligned} GQFEVD_{lk,h} &= \sum_{j=0}^{h-1} (\mathbf{e}_l^\top \mathbf{J}\mathbf{A}^j \mathbf{J}^\top \mathbf{J} \sqrt{v_i} \mathbf{D}\boldsymbol{\Omega}\mathbf{D}(\zeta_k \mathbf{D}_{kk})^{-1} \mathbf{e}_k)^2 \\ &\quad \times \left(\sum_{j=0}^{h-1} \sum_{k=1}^d (\mathbf{e}_l^\top \mathbf{J}\mathbf{A}^j \mathbf{J}^\top \mathbf{J} \sqrt{v_i} \mathbf{D}\boldsymbol{\Omega}\mathbf{D}(\zeta_k \mathbf{D}_{kk})^{-1} \mathbf{e}_k)^2 \right)^{-1}. \end{aligned}$$

Anche in questo caso, sfruttando le proprietà di posizione scala della distribuzione Laplace asimmetrica, possiamo riscrivere le *quantile forecast error variance decomposition* e *generalised quantile forecast error variance decomposition*, eliminando così il problema della stima di \mathbf{v} come:

$$QFEVD_{lk,h} = \sum_{j=0}^{h-1} (\mathbf{e}_l^\top \mathbf{J}\mathbf{A}^j \mathbf{J}^\top \mathbf{J} \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}} \mathbf{e}_k)^2 / \sum_{j=0}^{h-1} \sum_{k=1}^d (\mathbf{e}_l^\top \mathbf{J}\mathbf{A}^j \mathbf{J}^\top \mathbf{J} \mathbf{D}\boldsymbol{\Omega}^{\frac{1}{2}} \mathbf{e}_k)^2.$$

Analogamente si può ottenere una stima della *generalised quantile forecast error variance decomposition* che non dipenda da \mathbf{v} definita come:

$$GQFEVD_{lk,h} = \sum_{j=0}^{h-1} (\mathbf{e}_l^T \mathbf{J} \mathbf{A}^j \mathbf{J}^T \mathbf{J} \mathbf{D} \boldsymbol{\Omega} \mathbf{D} (\zeta_k \mathbf{D}_{kk})^{-1} \mathbf{e}_k)^2 \\ \times \left(\sum_{j=0}^{h-1} \sum_{k=1}^d (\mathbf{e}_l^T \mathbf{J} \mathbf{A}^j \mathbf{J}^T \mathbf{J} \mathbf{D} \boldsymbol{\Omega} \mathbf{D} (\zeta_k \mathbf{D}_{kk})^{-1} \mathbf{e}_k)^2 \right)^{-1}.$$

Capitolo 4

Algoritmi di stima

Lo scopo di questo capitolo è quello di presentare le tecniche di stima MCMC, spiegando brevemente il loro funzionamento, e mostrare come possono essere usate per la stima dei modelli introdotti nel precedente capitolo. Per ciascuno dei modelli presentati all'interno del capitolo 3 viene presentato il relativo algoritmo di stima. Nella parte finale del capitolo viene presentato un piccolo studio di simulazione, volto a confrontare le capacità degli algoritmi e dei modelli proposti. Viene riportato il codice per la stima dei modelli tramite gli algoritmi proposti e per la stima di analisi strutturali quantiliche all'interno dell' Appendice B.

4.1 Introduzione agli algoritmi MCMC

In ambito bayesiano la stima di quantità relative alle distribuzioni a posteriori, quando non note, risulta spesso non banale per via del numero elevato di parametri presenti e della complessità delle funzioni da massimizzare. L'utilizzo di tecniche *Markov Chain Monte Carlo* MCMC, permette di generare valori da una distribuzione di interesse e di stimare quantità e proprietà di interesse usando il campione generato, come fatto nei metodi *Monte Carlo*. Per quanto segue è stato preso a riferimento Andrieu et al. 2003. A scopo illustrativo definiamo θ parametro di interesse di cui si vuole conoscere la distribuzione $\pi(\theta)$. Per spiegare brevemente l'idea dietro alle tecniche MCMC, possiamo considerarle come degli strumenti che ci permettono, attraverso una

catena di Markov, di generare valori $\boldsymbol{\theta}^j \stackrel{asym}{\sim} \pi(\boldsymbol{\theta})$, che formeranno il nostro campione, e di muoverci nello spazio di $\boldsymbol{\theta}$. I valori generati $\boldsymbol{\theta}^j$ devono essere tali che, al crescere di j , la distribuzione dei $\boldsymbol{\theta}^j$ converga alla distribuzione $\pi(\boldsymbol{\theta})$. Vengono presentate due delle tecniche MCMC maggiormente utilizzate, e che verranno utilizzate anche negli algoritmi presentati successivamente. L'algoritmo *Metropolis Hastings* è il più importante e più popolare metodo MCMC, grazie anche al fatto che altri metodi, quali il *Gibbs Sampler*, che verrà presentato successivamente, possono essere visti come casi particolari del *Metropolis Hastings*. Data $\pi(\boldsymbol{\theta})$, densità da cui vogliamo generare, definiamo una densità accessoria, definita densità di transizione o candidata $q(\boldsymbol{\theta}^*|\boldsymbol{\theta}^j)$, che verrà usata, ad ogni step j dell'algoritmo, per generare dei valori candidati $\boldsymbol{\theta}^*$. Viene brevemente presentato l'algoritmo base di generazione tramite *Metropolis Hastings* nell'algoritmo 1.

Algorithm 1 Algoritmo Metropolis Hastings.

```

assegno valore iniziale  $\boldsymbol{\theta}^1$ 
for  $j = 2, \dots, n$  do
  genero  $\boldsymbol{\theta}^*$  da  $q(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{j-1})$ 
  calcolo  $\alpha(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{j-1}) = \min(1, \frac{\pi(\boldsymbol{\theta}^*)q(\boldsymbol{\theta}^{j-1}|\boldsymbol{\theta}^*)}{\pi(\boldsymbol{\theta}^{j-1})q(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{j-1})})$ 
  if  $u \sim \mathbf{U}(0, 1) \leq \alpha(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{j-1})$  then
     $\boldsymbol{\theta}^j \leftarrow \boldsymbol{\theta}^*$ 
  else
     $\boldsymbol{\theta}^j \leftarrow \boldsymbol{\theta}^{j-1}$ 
  end if
end for

```

Per creare un algoritmo *Metropolis Hastings* funzionante ed efficiente la scelta di $\boldsymbol{\theta}^1$ e di $q(\cdot|\cdot)$ risulta importante e può fortemente influenzare le capacità dell'algoritmo; in particolare $q(\cdot|\cdot)$ deve essere tale da permettere di raggiungere tutto lo spazio di $\boldsymbol{\theta}$ e di non rimanere bloccata in *loop*. Un caso particolare del *Metropolis Hastings*, detto *independence sampler*, si ha nel caso in cui la distribuzione $q(\cdot|\cdot)$ è simmetrica, ossia $q(\boldsymbol{\theta}^{j-1}|\boldsymbol{\theta}^j) = q(\boldsymbol{\theta}^j|\boldsymbol{\theta}^{j-1})$; in questo caso $\alpha(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{j-1}) = \min(1, \frac{\pi(\boldsymbol{\theta}^*)q(\boldsymbol{\theta}^{j-1}|\boldsymbol{\theta}^*)}{\pi(\boldsymbol{\theta}^{j-1})q(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{j-1})})$ si semplifica in $\alpha(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{j-1}) = \min(1, \frac{\pi(\boldsymbol{\theta}^*)}{\pi(\boldsymbol{\theta}^{j-1})})$.

Un caso particolare dell'algoritmo *Metropolis Hastings* che verrà principalmente utilizzato nel seguito del capitolo è il *Gibbs Sampler*. Immaginiamo di poter dividere il nostro parametro $\boldsymbol{\theta}$ in due o più gruppi di sotto-parametri, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_l)$. Scegliendo come densità di transizione $q(\boldsymbol{\theta}^* | \boldsymbol{\theta}^{j-1}) = \pi(\theta_i | \boldsymbol{\theta}_{-i}^{j-1})$, per $i = 1, \dots, l$, si ha che la probabilità α di accettazione si semplifica come:

$$\begin{aligned} \alpha(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{j-1}) &= \min \left(1, \frac{\pi(\boldsymbol{\theta}^*)q(\boldsymbol{\theta}^{j-1} | \boldsymbol{\theta}^*)}{\pi(\boldsymbol{\theta}^{j-1})q(\boldsymbol{\theta}^* | \boldsymbol{\theta}^{j-1})} \right) \\ &\quad \min \left(1, \frac{\pi(\boldsymbol{\theta}^*)\pi(\theta_i^{j-1} | \boldsymbol{\theta}_{-i}^{j-1})}{\pi(\boldsymbol{\theta}^{j-1})\pi(\theta_i^* | \boldsymbol{\theta}_{-i}^*)} \right) \\ &\quad \min \left(1, \frac{\pi(\boldsymbol{\theta}_{-i}^*)}{\pi(\boldsymbol{\theta}_{-i}^{j-1})} \right) = 1, \end{aligned}$$

portandoci quindi ad accettare sempre i valori proposti. La generazione tramite *Gibbs Sampler*, quando si avrà distribuzioni $\pi(\theta_i | \boldsymbol{\theta}_{-i}^{j-1})$ note, risulterà quindi preferibile. Un importante elemento da notare è che le tecniche MCMC, a differenza di altri metodi, non generano valori indipendenti né generano valori da $\pi(\boldsymbol{\theta})$, ma da una catena di valori correlati che converge alla distribuzione di interesse $\pi(\boldsymbol{\theta})$. Per permettere alla catena di approssimare al meglio $\pi(\boldsymbol{\theta})$, viene eliminata una prima parte dei valori generati, in cui la catena dei valori generati non si ritiene essere sufficiente vicina alla vera distribuzione $\pi(\boldsymbol{\theta})$, detta di *burn-in*, e si utilizzano per le procedure di inferenza solamente i valori successivi. Essendo inoltre i valori correlati tra loro, per avvicinarsi il più possibile ad un campione ideale di valori indipendenti può risultare utile effettuare un campionamento nella catena generata, tenendo solo un sottocampione costruito scegliendo generazioni ad intervalli regolare.

4.2 Algoritmi di stima

In questa sezione vengono presentati gli algoritmi di stima proposti per i modelli presentati nel corso del capitolo 3. In questa sezione si utilizzerà spesso, invece che \mathbf{D} , matrice diagonale di deviazione standard, la forma $\mathbf{d} = \text{diag}(\mathbf{D})$.

4.2.1 Algoritmo di stima modello VAR quantilico

Per la stima del modello VAR(p) presentato in 3.2 propongo di utilizzare un algoritmo MCMC, utilizzando dei passi di *Gibbs Sampling* per $(\boldsymbol{\beta}, \boldsymbol{\Omega}^{i-1}, \mathbf{v}^{i-1})$, date le densità a posteriori note derivate in 3.2, e di utilizzare un passo di *Metropolis Hastings* per la generazione di $\mathbf{d} = \text{diag}(\mathbf{D})$. Come densità di transizione per \mathbf{d} propongo di utilizzare un *random walk* uniforme, centrato sulla precedente iterazione. Nella stima di \mathbf{d} tramite *Metropolis Hastings* è necessario calcolare $\alpha_{MH_{\mathbf{d}}}$, probabilità di accettazione del valore proposto per \mathbf{d} , definita come:

$$\alpha_{MH_{\mathbf{d}}} = \alpha_{MH_{\mathbf{d}}}(\mathbf{d}^{i-1}, \mathbf{d}^*) = \frac{\pi(\mathbf{d}^* | \mathbf{y}, \mathbf{Z}, \boldsymbol{\beta}^{i-1}, \boldsymbol{\Omega}^{i-1}, \mathbf{v}^{i-1})}{\pi(\mathbf{d}^{i-1} | \mathbf{y}, \mathbf{Z}, \boldsymbol{\beta}^{i-1}, \boldsymbol{\Omega}^{i-1}, \mathbf{v}^{i-1})}. \quad (4.1)$$

Per l'utilizzo di un algoritmo *Gibbs Sampler* è necessario definire delle distribuzioni condizionate per i parametri; di seguito vengono riportate le distribuzioni a posteriori utilizzate per la generazione dei parametri. Come posteriori per $\boldsymbol{\beta}$ si assume $\boldsymbol{\beta} | \dots \sim \mathbf{N}(\widehat{\boldsymbol{\mu}}_{\boldsymbol{\beta}}^p, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}}^p)$, con

$$\begin{aligned} \widehat{\boldsymbol{\mu}}_{\boldsymbol{\beta}}^p &= \text{vec}((\mathbf{Y} - \mathbf{v} \otimes \boldsymbol{\lambda}^T \mathbf{D})^T \mathbf{V}^{-1} \mathbf{Z}^T (\mathbf{Z} \mathbf{V}^{-1} \mathbf{Z}^T)^{-1}) \\ \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}}^p &= ((\mathbf{Z} \mathbf{V}^{-1} \mathbf{Z}^T)^{-1}) \otimes \mathbf{D} \boldsymbol{\Omega} \mathbf{D}. \end{aligned}$$

Per \mathbf{v} , fattore latente, si assume la seguente distribuzione sui singoli elementi condizionati: $v_i | \dots \sim \text{Gig}(\frac{2-d}{2}, 2 + \boldsymbol{\lambda}^T \boldsymbol{\Omega}^{-1} \boldsymbol{\lambda}, \boldsymbol{\varepsilon}_i^T (\mathbf{D} \boldsymbol{\Omega} \mathbf{D})^{-1} \boldsymbol{\varepsilon}_i)$, con $\boldsymbol{\varepsilon}_i = \mathbf{y}_i - \mathbf{Z}_i \mathbf{B}$. Per $\boldsymbol{\Omega}$ viene assunta la seguente posteriori: $\boldsymbol{\Omega} | \mathbf{v}, \boldsymbol{\beta}, \mathbf{D}, \mathbf{Y}, \mathbf{X} \sim \text{IW}(\widehat{\mathbf{C}}_{\boldsymbol{\Omega}}^p, \widehat{c}^p)$, con:

$$\begin{aligned} \widehat{\mathbf{C}}_{\boldsymbol{\Omega}}^p &= \mathbf{C}_{\boldsymbol{\Omega}} + \mathbf{D}^{-1} (\mathbf{Y} - \mathbf{Z}^T \mathbf{B} - \mathbf{v} \otimes \boldsymbol{\lambda}^T \mathbf{D})^T \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{Z}^T \mathbf{B} - \mathbf{v} \otimes \boldsymbol{\lambda}^T \mathbf{D}) \mathbf{D}^{-1} \\ \widehat{c}^p &= c + n. \end{aligned}$$

Per generare \mathbf{d} si assume per $\pi(\mathbf{D} | \mathbf{y}, \mathbf{Z}, \boldsymbol{\beta}, \boldsymbol{\Omega}, \mathbf{v})$ una distribuzione proporzionale a:

$$\begin{aligned} \pi(\mathbf{D} | \mathbf{v}, \boldsymbol{\beta}, \boldsymbol{\Omega}, \mathbf{Y}, \mathbf{X}) &\propto |\mathbf{D}|^{-n} e^{-\frac{1}{2} (\mathbf{y} - \mathbf{X} \boldsymbol{\beta})^T (\mathbf{V} \otimes \mathbf{D} \boldsymbol{\Omega} \mathbf{D})^{-1} (\mathbf{y} - \mathbf{X} \boldsymbol{\beta})} \\ &\quad \times e^{(\mathbf{y} - \mathbf{X} \boldsymbol{\beta})^T (\mathbf{1}_n \otimes (\mathbf{D} \boldsymbol{\Omega}))^{-1} \boldsymbol{\lambda}}. \end{aligned}$$

L'algoritmo di stima viene riportato nell'algoritmo 2.

Algorithm 2 Algoritmo stima modello VAR quantilico.

inizializzo generando dalle rispettive priori $\mathbf{v}^1, \boldsymbol{\Omega}^1, \mathbf{d}^1, \boldsymbol{\beta}^1$

for $i = 2, \dots, n$ **do**

 genero \mathbf{v}^i da $\pi(\mathbf{v}|\boldsymbol{\Omega}^{i-1}, \boldsymbol{\beta}^{i-1}, \mathbf{d}^{i-1})$

 genero $\boldsymbol{\beta}^i$ da $\pi(\boldsymbol{\beta}|\boldsymbol{\Omega}^{i-1}, \mathbf{v}^i, \mathbf{d}^{i-1})$

 genero $\boldsymbol{\Omega}^*$ da $\pi(\boldsymbol{\Omega}|\boldsymbol{\beta}^{i-1}, \mathbf{v}^i, \mathbf{d}^{i-1})$.

$\mathbf{R}^i \leftarrow \mathbf{S}_\tau^{-1} \boldsymbol{\Omega}^* \mathbf{S}_\tau^{-1}$

$\boldsymbol{\Omega}^i \leftarrow \mathbf{S}_\tau \mathbf{R}^i \mathbf{S}_\tau$.

 genero \mathbf{d}^* $\leftarrow \mathbf{d}^{i-1} + \mathbf{c} \mathbf{U}_d(0, 1)$.

 calcolo α_{MH_d} come in (4.3)

if $u \sim \mathbf{U}(0, 1) \leq \alpha_{MH_d}$ **then**

$\mathbf{d}^i \leftarrow \mathbf{d}^*$

else

$\mathbf{d}^i \leftarrow \mathbf{d}^{i-1}$

end if

end for

4.2.2 Algoritmo per generazione unica della matrice di covarianza

In questa sezione viene presentato l'algoritmo di stima relativo al modello proposto in 3.2.1. Rispetto al modello precedente i parametri da stimare sono cambiati da $(\boldsymbol{\Omega}, D, \boldsymbol{\beta}, \mathbf{v})$ a $(\boldsymbol{\Omega}_d, \boldsymbol{\beta}, \mathbf{v})$, con $\boldsymbol{\Omega}_d = \boldsymbol{\zeta} \mathbf{D} \mathbf{R} \mathbf{D} \boldsymbol{\zeta}$. Per definire l'algoritmo di stima devo definire la probabilità di accettazione per i valori candidati di $\boldsymbol{\Omega}_d$, $\alpha_{MH_{\boldsymbol{\Omega}}}$, come:

$$\alpha_{MH_{\boldsymbol{\Omega}}}(\boldsymbol{\Omega}_d^{i-1}, \boldsymbol{\Omega}_d^*) = \frac{\pi(\boldsymbol{\Omega}_d^* | \mathbf{y}, \mathbf{Z}, \boldsymbol{\beta}^i, \mathbf{v}^1) \pi(\boldsymbol{\Omega}_d^{i-1} | \boldsymbol{\Omega}_d^*)}{\pi(\boldsymbol{\Omega}_d^{i-1} | \mathbf{y}, \mathbf{Z}, \boldsymbol{\beta}^i, \mathbf{v}^i) \pi(\boldsymbol{\Omega}_d^* | \boldsymbol{\Omega}_d^{i-1})}, \quad (4.2)$$

con densità di transizione scelta come:

$$\pi(\boldsymbol{\Omega}_d^1 | \boldsymbol{\Omega}_d^2) = \text{IW}(\boldsymbol{\Omega}_d^1, n \boldsymbol{\Omega}_d^2),$$

e $\pi(\boldsymbol{\Omega}_d | \mathbf{y}, \mathbf{Z}, \boldsymbol{\beta}, \mathbf{v})$ proporzionale a :

$$\pi(\boldsymbol{\Omega}_d | \mathbf{y}, \mathbf{Z}, \boldsymbol{\beta}, \mathbf{v}) \propto |\boldsymbol{\Omega}_d|^{-\frac{n}{2}} \mathbf{e}^{-\frac{1}{2} \text{Tr}(\mathbf{S}_\tau^{-1} (\mathbf{Y} - \mathbf{Z} \boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D} \boldsymbol{\lambda})^\top \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{Z} \boldsymbol{\beta} - \mathbf{v} \otimes \mathbf{D} \boldsymbol{\lambda}) \mathbf{S}_\tau^{-1} (\boldsymbol{\Omega}_d)^{-1})}.$$

Ricavando da Ω_d le matrici \mathbf{D} e Ω si ha come posteriori per (β, \mathbf{v}) : $(\beta \sim \mathbf{N}(\hat{\mu}_\beta^p, \hat{\Sigma}_\beta^p)$, con:

$$\begin{aligned}\hat{\mu}_\beta^p &= \text{vec}((\mathbf{Y} - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})^\top \mathbf{V}^{-1} \mathbf{Z}^\top (\mathbf{Z} \mathbf{V}^{-1} \mathbf{Z}^\top)^{-1}) \\ \hat{\Sigma}_\beta^p &= ((\mathbf{Z} \mathbf{V}^{-1} \mathbf{Z}^\top)^{-1}) \otimes \mathbf{D} \Omega \mathbf{D},\end{aligned}$$

e $v_i | \beta \Omega_d \sim \text{Gig}(\frac{2-d}{2}, 2 + \boldsymbol{\lambda}^\top \Omega^{-1} \boldsymbol{\lambda}, \boldsymbol{\varepsilon}_i^\top (\mathbf{D} \Omega \mathbf{D})^{-1} \boldsymbol{\varepsilon}_i)$, $\boldsymbol{\varepsilon}_i = \mathbf{y}_i - \mathbf{Z}_i \mathbf{B}$. L'algoritmo di stima, date le distribuzioni appena definite, viene riportato nell'algoritmo 3.

Algorithm 3 Algoritmo generazione Ω_d .

```

inizializzo generando dalla priori  $\mathbf{v}^1, \Omega_d^1, \beta^1$ 
for  $i = 2, \dots, n$  do
  genero  $\mathbf{v}^i$  da  $\pi(\mathbf{v} | \Omega_d^{i-1}, \beta^{i-1})$ 
  genero  $\beta^i$  da  $\pi(\beta | \Omega_d^{i-1}, \mathbf{v}^i)$ 
  genero  $\Omega_d^*$  da  $IW(n, n\Omega_d^{i-1})$ .
  calcolo  $\alpha_{MH\Omega}$  come in (4.2)
  if  $u \sim \mathbf{U}(0, 1) \leq \alpha_{MH\Omega}$  then
     $\Omega_d^i \leftarrow \Omega_d^*$ 
  else
     $\Omega_d^i \leftarrow \Omega_d^{i-1}$ 
  end if
end for

```

4.2.3 Algoritmo di stima VAR con restrizioni

Per la stima del modello VAR con restrizioni, presentato in 2.2, propongo di ricorrere, similmente a quanto fatto nelle precedenti sezioni, ad un algoritmo MCMC, sfruttando per la generazione di $(\delta, \Omega, \mathbf{v})$ l'algoritmo di *Gibbs Sampling*, e *Metropolis Hastings* per il parametro \mathbf{d} . Per stima di \mathbf{d} tramite *Metropolis Hastings* calcolo α_{MH_d} come:

$$\alpha_{MH_d}(\mathbf{d}^{i-1}, \mathbf{d}^*) = \frac{\pi(\mathbf{d}^* | \mathbf{y}, \mathbf{Z}, \boldsymbol{\delta}^{i-1}, \Omega^{i-1}, \mathbf{v}^{i-1})}{\pi(\mathbf{d}^{i-1} | \mathbf{y}, \mathbf{Z}, \boldsymbol{\delta}^{i-1}, \Omega^{i-1}, \mathbf{v}^{i-1})}. \quad (4.3)$$

Vengono qui riportate le priori utilizzate. Si utilizza come posteriori per $\boldsymbol{\delta}$ $\boldsymbol{\delta} | \dots \sim \mathbf{N}(\widehat{\boldsymbol{\mu}}_{\boldsymbol{\delta}}^p, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\delta}}^p)$:

$$\begin{aligned}\widehat{\boldsymbol{\mu}}_{\boldsymbol{\delta}}^p &= \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\delta}}^p (\mathbf{C}^\top \text{vec}((\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1}(\mathbf{Y}^* - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})^\top \mathbf{V}^{-1} \mathbf{Z}^\top) + \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1} \boldsymbol{\mu}_{\boldsymbol{\delta}}) \\ \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\delta}}^p &= (\boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1} + \mathbf{C}^\top (\mathbf{Z}\mathbf{V}^{-1} \mathbf{Z}^\top) \otimes (\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} \mathbf{C})^{-1},\end{aligned}$$

per \mathbf{v} si ha la seguente distribuzione per il singolo elemento, $v_i, i = 1, \dots, n$, $v_i | \boldsymbol{\Omega}, \boldsymbol{\delta}, \mathbf{D}, \mathbf{Y}, \mathbf{X} \sim \text{Gig}(\frac{2-d}{2}, 2 + \boldsymbol{\lambda}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{\lambda}, \boldsymbol{\varepsilon}_i^{*\top} (\mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} \boldsymbol{\varepsilon}_i^*)$, con $\boldsymbol{\varepsilon}_i^* = \mathbf{Y}_i^* - \text{invec}(\mathbf{X}\mathbf{C}\boldsymbol{\delta})_i$. Per $\boldsymbol{\Omega}$ si ha come distribuzione a posteriori, $\boldsymbol{\Omega} | \mathbf{v}, \boldsymbol{\delta}, \mathbf{D}, \mathbf{Y}, \mathbf{X} \sim \text{IW}(\widehat{\mathbf{C}}_{\boldsymbol{\Omega}}^p, \widehat{c}^p)$, con:

$$\begin{aligned}\widehat{\mathbf{C}}_{\boldsymbol{\Omega}}^p &= \mathbf{C}_{\boldsymbol{\Omega}} + \mathbf{D}^{-1} (\mathbf{Y}^* - \text{invec}(\mathbf{X}\mathbf{C}\boldsymbol{\delta}) - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})^\top \mathbf{V}^{-1} (\mathbf{Y}^* - \text{invec}(\mathbf{X}\mathbf{C}\boldsymbol{\delta}) - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D}) \mathbf{D}^{-1} \\ \widehat{c}^p &= c + n.\end{aligned}$$

Come distribuzione a posteriori per \mathbf{D} si assume:

$$\begin{aligned}\pi(\mathbf{D} | \mathbf{v}, \boldsymbol{\delta}, \boldsymbol{\Omega}, \mathbf{Y}, \mathbf{X}) &\propto |\mathbf{D}|^{-n} \\ &\times e^{-\frac{1}{2} (\mathbf{y}^* - \mathbf{X}\mathbf{C}\boldsymbol{\delta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})^\top (\mathbf{V} \otimes \mathbf{D}\boldsymbol{\Omega}\mathbf{D})^{-1} (\mathbf{y}^* - \mathbf{X}\mathbf{C}\boldsymbol{\delta} - \mathbf{v} \otimes \mathbf{D}\boldsymbol{\lambda})}.\end{aligned}$$

L'algoritmo di stima per un modello VAR con restrizioni viene riportato nell'algoritmo 4.

4.2.4 Algoritmo di stima *spike-and-slab*

In questa sezione viene presentato l'algoritmo di stima proposto per il modello VAR(p) quantilico bayesiano con priori *spike-and-slab*. Vengono qui riportate le distribuzioni usate nell'algoritmo:

$v_i | \dots \sim \text{GIG}(\frac{2-d}{2}, 2 + \boldsymbol{\lambda}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{\lambda}, \boldsymbol{\varepsilon}_i^\top (\mathbf{D}\mathbf{S}_\tau \mathbf{R} \mathbf{S}_\tau \mathbf{D})^{-1} \boldsymbol{\varepsilon}_i)$, con $\boldsymbol{\varepsilon}_i = \mathbf{y}_i - \mathbf{m}\mathbf{u} - \mathbf{Z}_i \mathbf{B}$. Per $\boldsymbol{\mu}$ si utilizza come posteriori $\boldsymbol{\mu} | \dots \sim \mathbf{N}(\boldsymbol{\mu}^p, \boldsymbol{\Sigma}^p)$, con:

$$\begin{aligned}\boldsymbol{\mu}^p &= \frac{(\mathbf{v}^{-1} (\mathbf{Y} - \mathbf{Z}\mathbf{B} - \mathbf{v}\mathbf{D}\boldsymbol{\lambda}))}{\sum_{i=1}^n \frac{1}{v_i}} \\ \boldsymbol{\Sigma}^p &= \left(\sum_{i=1}^n \frac{1}{v_i} \right) \mathbf{D}\mathbf{S}_\tau \boldsymbol{\Omega} \mathbf{S}_\tau \mathbf{D}.\end{aligned}$$

Per il parametro $\boldsymbol{\Omega}$ si ha $\boldsymbol{\Omega} | \dots \sim \text{IW}(\widehat{\mathbf{C}}_{\boldsymbol{\Omega}}^p, \widehat{c}^p)$, con:

$$\begin{aligned}\widehat{\mathbf{C}}_{\boldsymbol{\Omega}}^p &= \mathbf{C}_{\boldsymbol{\Omega}} + \mathbf{D}^{-1} (\mathbf{Y} - \mathbf{1}_n \boldsymbol{\mu} - \mathbf{Z}^\top \mathbf{B} - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})^\top \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{1}_n \boldsymbol{\mu} - \mathbf{Z}^\top \mathbf{B} - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D}) \mathbf{D}^{-1} \\ \widehat{c}^p &= c + n.\end{aligned}$$

Algorithm 4 Algoritmo VAR con restrizioni.

inizializzo generando dalla priori $\mathbf{v}^1, \boldsymbol{\Omega}^1, \mathbf{d}^1, \boldsymbol{\delta}^1$

for $i = 2, \dots, n$ **do**

genero \mathbf{v}^i da $\pi(\mathbf{v} | \boldsymbol{\Omega}^{i-1}, \boldsymbol{\delta}^{i-1}, \mathbf{d}^{i-1})$

genero $\boldsymbol{\delta}^i$ da $\pi(\boldsymbol{\delta} | \boldsymbol{\Omega}^{i-1}, \mathbf{v}^i, \mathbf{d}^{i-1})$

genero $\boldsymbol{\Omega}^*$ da $\pi(\boldsymbol{\Omega} | \boldsymbol{\delta}^i, \mathbf{v}^i, \mathbf{d}^{i-1})$.

$\mathbf{R}^i \leftarrow \mathbf{S}_\tau^{-1} \boldsymbol{\Omega}^* \mathbf{S}_\tau^{-1}$

$\boldsymbol{\Omega}^i \leftarrow \mathbf{S}_\tau \mathbf{R}^i \mathbf{S}_\tau$.

$\mathbf{d}^* \leftarrow \mathbf{d}^{i-1} + \mathbf{c} \mathbf{U}_d(0, 1)$

calcolo α_{MH_d} come in (4.3)

if $u \sim \mathbf{U}(0, 1) \leq \alpha_{MH_d}$ **then**

$\mathbf{d}^i \leftarrow \mathbf{d}^*$

else

$\mathbf{d}^i \leftarrow \mathbf{d}^{i-1}$

end if

end for

mentre per il parametro $\boldsymbol{\beta}$ si ha: $\boldsymbol{\beta} | \dots \sim \mathbf{N}(\widehat{\boldsymbol{\mu}}_\beta^p, \widehat{\boldsymbol{\Sigma}}_\beta^p)$, con:

$$\begin{aligned} \widehat{\boldsymbol{\mu}}_\beta^p &= \widehat{\boldsymbol{\Sigma}}_\beta^p (\text{vec}(\boldsymbol{\Omega}^{-1}(\mathbf{Y} - \mathbf{v} \otimes \boldsymbol{\lambda}^\top \mathbf{D})^\top \mathbf{V}^{-1} \mathbf{Z}^\top) + \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\mu}_\beta) \\ \widehat{\boldsymbol{\Sigma}}_\beta^p &= (\boldsymbol{\Sigma}_\beta^{-1} + (\mathbf{Z} \mathbf{V}^{-1} \mathbf{Z}^\top) \otimes \boldsymbol{\Omega}^{-1})^{-1}. \end{aligned}$$

La posteriori per k_{l1} , parametro di scala relativo alle righe del modello, risulta essere : $k_{l1} | \dots \sim \text{IG}(a_k + \frac{M}{2}, b_k + \frac{1}{2} \sum_{j=1}^M \frac{\mathbf{B}_{lk}^2}{\sigma_{\varepsilon_j} ((1-\gamma_{lk})k_0 + \gamma_{lk})})$ per $l = 1, \dots, d$. La posteriori per σ_{ε_j} , parametro di scala relativo alle colonne del modello, utilizzata è: $\sigma_{\varepsilon_j} | \dots \sim \text{IG}(a_\varepsilon + \frac{d}{2}, \sum_{l=1}^d \frac{\mathbf{B}_{lj}^2}{2k_{l1}(k_0(1-\gamma_{lj}) + \gamma_{lj})} + b_\varepsilon)$ per $j = 1, \dots, M$. La distribuzione a posteriori di π_{0l} , probabilità di generare un elemento dalla *slab* risulta $\pi_{0l} | \dots \sim \text{Beta}(a_\pi + \sum_{j=1}^M \gamma_{lj}, b_{\pi_0} + \sum_{j=1}^M (1-\gamma_{lj}))$, per $l = 1, \dots, d$. La distribuzione a posteriori per γ_{lj} risulta una binomiale con probabilità di successo proporzionali a:

$$\begin{aligned} \pi(\gamma_{lj} = 1 | \dots) &\propto \frac{\pi_{0l}}{\sqrt{k_{l1} \sigma_{\varepsilon_j}}} \mathbf{e}^{-\frac{\mathbf{B}_{lj}^2}{2k_{l1} \sigma_{\varepsilon_j}}} \\ \pi(\gamma_{lj} = 0 | \dots) &\propto \frac{1 - \pi_{0l}}{\sqrt{k_0 k_{l1} \sigma_{\varepsilon_j}}} \mathbf{e}^{-\frac{\mathbf{B}_{lj}^2}{2k_0 k_{l1} \sigma_{\varepsilon_j}}}. \end{aligned}$$

Per quanto riguarda stima e generazione di \mathbf{D} vengono usate le stesse quantità definite in 4.2.1. L'algoritmo di stima per un modello VAR quantilico con priori *spike-and-slab* viene riportato nell'algoritmo 5.

Algorithm 5 Algoritmo *spike-and-slab*.

inizializzo generando dalla priori $(\mathbf{d}^0, \boldsymbol{\beta}^0, \boldsymbol{\gamma}^0, \boldsymbol{\pi}_0^0, \boldsymbol{\mu}^0, \boldsymbol{\Omega}^0, \mathbf{v}^0, K_1^0)$

for $i = 2, \dots, n$ **do**

 genero \mathbf{v}^i da $\pi(\mathbf{v}|\boldsymbol{\Omega}^{i-1}, \boldsymbol{\beta}^{i-1}, \mathbf{d}^{i-1}, \boldsymbol{\gamma}^{i-1}, \boldsymbol{\mu}^{i-1}, \boldsymbol{\pi}_0^{i-1}, \mathbf{k}_1^{i-1})$.

 genero $\boldsymbol{\pi}_0$ da $\pi(\boldsymbol{\pi}_0|\boldsymbol{\Omega}^{i-1}, \boldsymbol{\beta}^{i-1}, \mathbf{d}^{i-1}, \boldsymbol{\gamma}^{i-1}, \boldsymbol{\mu}^{i-1}, \mathbf{v}^i, \mathbf{k}_1^{i-1})$.

 genero $\boldsymbol{\gamma}$ da $\pi(\boldsymbol{\gamma}|\boldsymbol{\Omega}^{i-1}, \boldsymbol{\beta}^{i-1}, \mathbf{d}^{i-1}, \boldsymbol{\mu}^{i-1}, \boldsymbol{\pi}_0^i, \mathbf{v}^i, \mathbf{k}_1^{i-1})$.

 genero \mathbf{k}_1 da $\pi(\mathbf{k}_1|\boldsymbol{\Omega}^{i-1}, \boldsymbol{\beta}^{i-1}, \mathbf{d}^{i-1}, \boldsymbol{\mu}^{i-1}, \boldsymbol{\pi}_0^i, \mathbf{v}^i, \boldsymbol{\gamma}^i)$.

 genero $\boldsymbol{\beta}^i$ da $\pi(\boldsymbol{\beta}|\boldsymbol{\Omega}^{i-1}, \mathbf{d}^{i-1}, \boldsymbol{\mu}^{i-1}, \boldsymbol{\pi}_0^i, \mathbf{v}^i, \boldsymbol{\gamma}^i, \mathbf{k}_1^i)$.

 genero $\boldsymbol{\mu}^i$ da $\pi(\boldsymbol{\mu}|\boldsymbol{\Omega}^{i-1}, \mathbf{d}^{i-1}, \boldsymbol{\beta}^i, \boldsymbol{\pi}_0^i, \mathbf{v}^i, \boldsymbol{\gamma}^i, \mathbf{k}_1^i)$.

 genero $\boldsymbol{\Omega}^*$ da $\pi(\boldsymbol{\Omega}|\boldsymbol{\mu}^i, \mathbf{d}^{i-1}, \boldsymbol{\beta}^i, \boldsymbol{\pi}_0^i, \mathbf{v}^i, \boldsymbol{\gamma}^i, \mathbf{k}_1^i)$.

$\mathbf{R}^i \leftarrow \mathbf{S}_\tau^{-1} \boldsymbol{\Omega}^* \mathbf{S}_\tau^{-1}$

$\boldsymbol{\Omega}^i \leftarrow \mathbf{S}_\tau \mathbf{R}^i \mathbf{S}_\tau$.

$\mathbf{d}^* \leftarrow \mathbf{d}^{i-1} + c\mathbf{U}_d(0, 1)$.

 calcolo α_{MH_d} come in (4.3)

if $u \sim \mathbf{U}(0, 1) \leq \alpha_{MH_d}$ **then**

$\mathbf{d}^i \leftarrow \mathbf{d}^*$

else

$\mathbf{d}^i \leftarrow \mathbf{d}^{i-1}$

end if

end for

4.3 Esempi con dati simulati

Per valutare le performance dei modelli e degli algoritmi introdotti è stato eseguito un breve studio di simulazione, valutando le performance degli algoritmi e modelli proposti con differenti ritardi p e dimensioni d .

4.3.1 Strumenti di valutazione

Vengono presentate in questa sezione le misure e i test che verranno utilizzati per valutare la qualità delle catene generate, e l'adeguatezza delle stime ottenute. Data la rapidità con cui il numero di parametri da stimare cresce, al crescere di d , risulta difficile utilizzare diagnostiche grafiche per valutare gli algoritmi; l'unico caso in cui si può ricorrere a diagnostiche grafiche è per gli elementi di \mathbf{D} , che sono di particolare interesse essendo \mathbf{D} l'unico parametro generato tramite *Metropolis Hastings*. Per valutare le performance dei modelli e confrontarli tra loro vengono utilizzati prevalentemente il *potential scale reduction factors* nella sua forma sintetica multivariata, e la *effective sample size*. Per quanto riguarda la valutazione delle capacità del modello di stimare la reale distribuzione nel quantile τ , vengono utilizzati i test LRuc e LRcc per valutare l'effettiva capacità di stima del quantile. Di seguito vengono brevemente presentate ed introdotte queste misure e test. Il *potential scale reduction factors* ci indica quanto, ipotizzando una serie MCMC che vada all'infinito, potremmo ridurre la scala della distribuzione generata. Quando il valore stimato del *potential scale reduction factors* si avvicina a 1 si avrà quindi una distribuzione vicina a quella possibile; nella pratica valori inferiori a 1.1/1.2 sono ritenuti accettabili. Si può considerare quindi il *potential scale reduction factors* come una misura della convergenza della catena generata alla distribuzione asintotica. Nel caso in cui si abbia un parametro multidimensionale, come nel nostro caso, invece che valutare il *potential scale reduction factors* per ogni sua singola componente si può utilizzare il *potential scale reduction factors* multivariato come misura di sintesi per l'insieme di parametri, calcolato come limite superiore dei *potential scale reduction factors* univariati per i parametri stimati. Generando tramite MCMC si generano valori correlati tra loro, questo porta ad avere una minore informazione

sulla distribuzione target rispetto a quanto si otterrebbe generando direttamente dalla distribuzione target; la *effective sample size* misura la bontà del campione generato, confrontandolo con un campione indipendente ed identicamente distribuito per la distribuzione target. Più la *effective sample size* del campione generato è vicina al numero di elementi generati, maggiore sarà l'informazione offerta, misurata in numero di estrazioni indipendenti equivalenti dalla distribuzione target. Per valutare la capacità dei modelli stimati di prevedere il comportamento della serie nei quantili modellati ho deciso di utilizzare il test LRcc, test di copertura condizionata, di Christoffersen 1998. Il test LRcc di viene definito come $LRcc = LRuc + LRind$, con LRind test di indipendenza e LRuc, test di copertura incondizionata di Kupiec 1995a. Il test LRuc di Kupiec 1995a può esser usato per confrontare la proporzione di elementi osservati che superano i quantili stimati dal modello con il quantile di riferimento. Data I_1, \dots, I_n sequenza di funzioni indicatrici, il test assume una verosimiglianza binomiale per la proporzione di elementi, funzioni indicatrici, $I_i = 1$. Il test viene definito come:

$$LRuc = 2 \log[L(\hat{\pi}, I_1, \dots, I_n) - L(\pi_0, I_1, \dots, I_n)] \stackrel{asym}{\sim} \chi^2(1),$$

con $\hat{\pi}$ proporzione campionaria di $I_i = 1$ e π_0 quantile/proporzione da testare come ipotesi nulla. Nel caso quantilico, il test viene utilizzato per valutare se la proporzione campionaria di elementi che superano il quantile τ è statisticamente diversa da τ . Il test di indipendenza LRind viene definito come test per valutare l'indipendenza degli elementi I_1, \dots, I_n ; l'ipotesi nulla equivale all'assunzione di una distribuzione binomiale per I_1, \dots, I_n , e viene testata con ipotesi alternativa che I_1, \dots, I_n sia generata da un processo di Markov di primo ordine. Il test è quindi definito come:

$$LRind = 2 \log[L(\hat{\pi}_{01}, \hat{\pi}_{11}, I_1, \dots, I_n) - L(\hat{\pi}, I_1, \dots, I_n)] \stackrel{asym}{\sim} \chi^2(1),$$

con $\hat{\pi}_{01}, \hat{\pi}_{11}$ probabilità stimate di transizione da uno stato all'altro della catena, e $\hat{\pi}$ probabilità stimata con l'assunzione di distribuzione binomiale. Trattandosi di test definiti tramite log-rapporto di verosimiglianza, si ha che la distribuzione asintotica del test LRcc=LRuc+LRind è $\chi^2(2)$. Nel caso quantilico il test LRcc viene utilizzato per valutare congiuntamente che la proporzione di elementi che superano il quantile non sia significativamente

diversa dal quantile di riferimento, e che non vi sia dipendenza di primo ordine tra gli elementi che superano il quantile.

4.3.2 Studio dei dati simulati

Come prima simulazione è stato generato un processo VAR(1) di lunghezza $n = 500$, con $d = 4$, dimensione di \mathbf{y} , e con 5 covariate. Ai dati simulati sono stati applicati gli algoritmi di stima introdotti nella precedente sezione, nei quantili $\tau = (0.5, 0.7, 0.9)$. Trattando una serie simulata volutamente simmetrica, il comportamento atteso nei quantili $(0.7, 0.9)$ è lo stesso che nei quantili $(0.3, 0.1)$, per cui è sufficiente, per valutare le performance di stima dell'algoritmo, stimare uno dei due. Il processo è stato generato come:

$$\begin{bmatrix} y_{i,1} \\ y_{i,2} \\ y_{i,3} \\ y_{i,4} \end{bmatrix} = \begin{bmatrix} 3.8 \\ 3.8 \\ 3 \\ 3 \end{bmatrix} + \begin{bmatrix} 0.40 & 0.10 & 0.10 & 0.10 \\ 0.10 & 0.40 & 0.10 & 0.10 \\ 0.10 & 0.10 & 0.40 & 0.10 \\ 0.10 & 0.10 & 0.10 & 0.40 \end{bmatrix} \begin{bmatrix} y_{i-1,1} \\ y_{i-1,2} \\ y_{i-1,3} \\ y_{i-1,4} \end{bmatrix} + \begin{bmatrix} 0.8 & 0.8 & 0.8 & 0.8 & 0.8 \\ 0.0 & 0.8 & 0.8 & 0.0 & 0.8 \\ 0.8 & 0.0 & 0.0 & 0.8 & 0.0 \\ 0.8 & 0.8 & 0.0 & 0.8 & 0.8 \end{bmatrix} \begin{bmatrix} x_{i,1} \\ x_{i,2} \\ x_{i,3} \\ x_{i,4} \\ x_{i,5} \end{bmatrix} + \boldsymbol{\varepsilon}_i.$$

Per la stima del processo tramite algoritmi MCMC sono state effettuate 20000 iterazioni, di cui 10000 di *burn-in*, tenendo poi nel campione solo uno stato della catene ogni 10, così da ridurre la dipendenza tra elementi successivi della catena. Di seguito viene riportata una tabella con la stima del test LRcc, implementato nel pacchetto GAS di R(David, Kris e Catania 2019), per valutare l'adeguatezza delle stime.

algoritmo	quantile	LRcc	p-value
2	0.5	0.532	0.766
2	0.7	0.001	0.969
2	0.9	2.195	0.138
3	0.5	1.532	0.766
3	0.7	0.035	0.885
3	0.9	1.742	0.186
4	0.5	0.062	0.487
4	0.7	1.438	0.564
4	0.9	3.366	0.186
5	0.5	0.484	0.781
5	0.7	0.000	0.999
5	0.9	9.271	0.002

Tabella 4.1: Valutazione tramite test LRcc delle capacità di stima dei quantili per le simulazioni con $d = 4$.

Tutti i modelli stimati risultano avere una buona capacità di stima dei rispettivi quantili di riferimento. Si può notare che, al crescere di τ , quantile di stima, si riduce la precisione delle stime in tutti i modelli utilizzati. Le peggiori performance all'avvicinarsi del quantile a 1 sono preventivabili, in quanto l'informazione contenuta nei dati sulla distribuzione nel quantile diminuisce più ci si avvicina alle code della distribuzione. Per poter valutare gli aspetti relativi alla qualità delle catene MCMC generate, sono riportati nelle tabelle 4.2,4.3 i valori del *potential scale reduction factors* multivariato e la *effective sample size* media per i parametri generati dalle catene.

algoritmo	quantile	mult. psrf	eff. sample size	algoritmo	quantile	mult. psrf	eff. sample size
2	0.5	1.02	793	2	0.5	1.01	4800
2	0.7	1.02	632	2	0.7	1.02	4578
2	0.9	1.08	327	2	0.9	1.07	1642
3	0.5	1.01	621	3	0.5	1.01	5041
3	0.7	1.03	587	3	0.7	1.01	4964
3	0.9	1.12	241	3	0.9	1.03	4403
4	0.5	1.05	691	4	0.5	1.01	4697
4	0.7	1.02	676	4	0.7	1.02	4518
4	0.9	1.09	327	4	0.9	1.05	2049
5	0.5	1.01	695	5	0.5	1.13	4778
5	0.7	1.03	554	5	0.7	1.19	2799
5	0.9	1.04	348	5	0.9	1.09	994

Tabella 4.2: Tabelle riassuntiva delle serie generate per \mathbf{D} e β , con dimensione $d = 4$.

algoritmo	quantile	mult.psrff	eff. sample size	parametro	quantile	mult.psrff	eff. sample size
2	0.5	1.01	4905	π_0	0.5	1.05	5046
2	0.7	1.01	2605	π_0	0.7	1.05	4437
2	0.9	1.04	560	π_0	0.9	1.05	3693
3	0.5	1.02	970	σ_ε	0.5	1.01	4940
3	0.7	1.02	805	σ_ε	0.7	1.01	4654
3	0.9	1.04	416	σ_ε	0.9	1.02	4418
4	0.5	1.01	4168	μ	0.5	1.08	1098
4	0.7	1.03	2503	μ	0.7	1.04	885
4	0.9	1.02	887	μ	0.9	1.09	261
5	0.5	1.00	4778	\mathbf{K}_1	0.5	1.03	4895
5	0.7	1.01	2799	\mathbf{K}_1	0.7	1.01	4593
5	0.9	1.06	531	\mathbf{K}_1	0.9	1.02	4660

Tabella 4.3: Tabelle riassuntiva delle serie generate per Σ e per i parametri relativi al modello con priori *spike-and-slab*, dimensione $d = 4$.

L'algoritmo 4.2.2, con la generazione congiunta di \mathbf{D} e Ω in uno *step* Metropolis-Hastings, risulta avere performance leggermente peggiori degli altri algoritmi in termini di *potential scale reduction factors* e *effective sample size* per la generazione di \mathbf{D} , mentre le performance per la generazione di Ω risultano nettamente peggiori; al contrario per la generazione di β sembra avere performance più elevate. Vengono riportati i grafici delle autocorrelazioni per i valori generati per \mathbf{D} per gli algoritmi 3, 2 in figura 4.1. L'algoritmo per la stima di un VAR con restrizioni, dovendo stimare meno parametri, risul-

ta avere una *effective sample size* migliore del corrispondente modello VAR quantilico. La generazione di $\boldsymbol{\mu}$ nell'algoritmo *spike-and-slab*, essendo strettamente collegata alla generazione di \mathbf{D} risulta poco efficiente se confrontata agli altri parametri. L'algoritmo *spike-and-slab*, seleziona correttamente le variabili diverse da zero all'interno della matrice dei coefficienti di regressione, attribuendo elevata probabilità agli elementi diversi da zero, e probabilità minore agli elementi nulli. A titolo esemplificativo viene riportata la tabella dei coefficienti di regressione e la stima delle probabilità di appartenere alla *slab* per $\tau = 0.9$.

0.90	0.93	0.91	0.90	0.89	0.80	0.80	0.80	0.80	0.80
0.26	0.88	0.66	0.35	0.84	-	0.80	0.80	-	0.80
0.92	0.39	0.39	0.94	0.41	0.80	-	-	0.80	-
0.96	0.87	0.41	0.93	0.89	0.80	0.80	-	0.80	0.80

Tabella 4.4: Confronto tra probabilità a posteriori di essere generato dalla *slab* e vero valore del parametro con cui si è generata la distribuzione.

Per valutare le performance degli algoritmi presentati in un caso leggermente più complesso è stato effettuato un secondo set di simulazioni, generando in questo caso da un processo VAR(2), con $d = 6$, $n = 500$ e numero di covariate $m = 2$. Il modello è stato generato come:

$$\begin{bmatrix} y_{i,1} \\ y_{i,2} \\ y_{i,3} \\ y_{i,4} \\ y_{i,5} \\ y_{i,6} \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \end{bmatrix} + \begin{bmatrix} 0.2 & 0.2 & 0 & 0.2 & 0 & 0.2 \\ 0 & 0.3 & 0 & 0.2 & 0 & 0.2 \\ 0.2 & 0.2 & 0.3 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0 & 0.2 & 0.2 & 0.2 & 0 \\ 0.0 & 0 & 0 & 0 & 0.2 & 0 \\ 0.2 & 0 & 0.2 & 0 & 0.2 & 0.2 \end{bmatrix} \begin{bmatrix} y_{i-1,1} \\ y_{i-1,2} \\ y_{i-1,3} \\ y_{i-1,4} \\ y_{i-1,5} \\ y_{i-1,6} \end{bmatrix} + \begin{bmatrix} -0.2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.2 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.2 \end{bmatrix} \begin{bmatrix} y_{i-2,1} \\ y_{i-2,2} \\ y_{i-2,3} \\ y_{i-2,4} \\ y_{i-2,5} \\ y_{i-2,6} \end{bmatrix} + \begin{bmatrix} 0.6 & 0.6 \\ 0.6 & 0.6 \\ 0.6 & 0.6 \\ 0.6 & 0.6 \\ 0.6 & 0.6 \\ 0.6 & 0.6 \end{bmatrix} \begin{bmatrix} x_{i,1} \\ x_{i,2} \end{bmatrix} + \boldsymbol{\varepsilon}_i,$$

con l'aggiunta di un disturbo Normale d -variato, $\boldsymbol{\varepsilon}_i$, a componenti dipendenti. Anche in questo caso, essendo il processo simmetrico, viene valutata la capacità di stima nei quantili (0.5, 0.7, 0.9). Vengono generate 5 catene MCMC in parallelo, con 10000 iterazioni di *burn-in*, tenendo poi una osservazione su 10 delle successive 10000 come parte del campione. Vengono riportati nella seguente tabella i risultati dei test LRcc per valutare l'adeguatezza delle stime effettuate.

algoritmo	quantile	LRcc	p-value
2	0.5	1.029	0.597
2	0.7	2.343	0.309
2	0.9	4.122	0.127
3	0.5	0.560	0.756
3	0.7	5.293	0.071
3	0.9	13.135	0.002
4	0.5	1.349	0.509
4	0.7	0.117	0.732
4	0.9	0.065	0.798
5	0.5	0.942	0.624
5	0.7	2.641	0.267
5	0.9	3.525	0.172

Tabella 4.5: Valutazione tramite test LRcc delle capacità di stima dei quantili per le simulazioni con $d = 6$.

Come era prevedibile l'algoritmo 4.2.2 con $d = 6$ non riesce a fornire una stima accettabile del quantile 0.9. L'algoritmo con restrizioni, sotto corretta specificazione delle restrizioni, stima meno parametri e risulta il più preciso. Viene confrontata l'autocorrelazione degli elementi di \mathbf{D} generati dall'algoritmo 4.2.2 e dall'algoritmo 4.2.1, figura 4.2. Si nota dai grafici come, per i valori generati dall'algoritmo 4.2.2, ci sia una maggiore autocorrelazione che permane a lungo, rendendo i valori generati per \mathbf{D} poco informativi e influenzando le performance generali del modello. L'algoritmo 4.2.4 e 4.2.1 ottengono performance simili. Rispetto alle simulazioni con $d = 4$ si ha un peggioramento generale della capacità di stima del modello VAR. Vengono riportate le analisi dei valori generati dalle catene in termini di *potential scale reduction factors* e *effective sample size* all'interno delle tabelle 4.6 e 4.7.

La qualità delle catene, specie per quanto riguarda \mathbf{D} , risulta nettamente peggiore rispetto al caso con $d = 4$. Generando congiuntamente tramite *step Metropolis-Hastings* gli elementi di \mathbf{D} si ha che al crescere della dimensione del parametro le stime risultano più correlate nel tempo, riducendo la *effective sample size* e peggiorando la qualità delle stime a parità di iterazioni.

algoritmo	quantile	mult.psrfr	eff. sample size	algoritmo	quantile	mult. psrfr	eff.sample size
2	0.5	1.04	442	2	0.5	1.03	4601
2	0.7	1.05	319	2	0.7	1.05	4149
2	0.9	1.20	151	2	0.9	1.09	1518
3	0.5	1.08	162	3	0.5	1.03	5026
3	0.7	1.29	128	3	0.7	1.04	4967
3	0.9	1.89	75	3	0.9	1.40	4119
4	0.5	1.07	506	4	0.5	1.05	4636
4	0.7	1.03	475	4	0.7	1.03	3957
4	0.9	1.21	154	4	0.9	1.21	1929
5	0.5	1.05	439	5	0.5	1.10	2607
5	0.7	1.09	390	5	0.7	1.17	2369
5	0.9	1.27	194	5	0.9	1.28	1217

Tabella 4.6: Tabelle riassuntiva delle serie generate per \mathbf{D} e β , dimensione $d = 6$.

Anche in questo caso la qualità delle catene generate per Ω e β decresce all'avvicinarsi di τ agli estremi $(0, 1)$. A parte per μ , strettamente legato a \mathbf{D} , il resto dei parametri aggiuntivi presenti nell'algoritmo 5 non presentano particolari variazioni nelle performance al crescere di d da 4 a 6. Il tasso di

algoritmo	quantile	mult.psrfr	eff. sample size	parametro	quantile	mult.psrfr	eff. sample size
2	0.5	1.01	4211	π_0	0.5	1.01	4901
2	0.7	1.01	2719	π_0	0.7	1.00	3739
2	0.9	1.04	661	π_0	0.9	1.00	3288
3	0.5	1.08	225	σ_ε	0.5	1.01	4847
3	0.7	1.12	180	σ_ε	0.7	1.02	4481
3	0.9	1.69	72	σ_ε	0.9	1.01	3961
4	0.5	1.00	4626	μ	0.5	1.03	909
4	0.7	1.00	2597	μ	0.7	1.09	535
4	0.9	1.07	830	μ	0.9	1.15	196
5	0.5	1.01	3120	\mathbf{K}_1	0.5	1.00	4839
5	0.7	1.03	1295	\mathbf{K}_1	0.7	1.00	4806
5	0.9	1.08	404	\mathbf{K}_1	0.9	1.01	4496

Tabella 4.7: Tabelle riassuntiva delle serie generate per Σ e per i parametri della *spike-and-slab*, dimensione $d = 6$.

corretta classificazione nella scelta dei parametri diversi da zero, utilizzando le medie dei valori generati per γ come posteriori della probabilità del singolo elemento e π per le soglie, uguali per ogni riga di \mathbf{B} , risulta pari a 0.91. Si è cercato di impostare un ulteriore set di simulazioni, per valutare se fosse

possibile utilizzare i metodi proposti con VAR di dimensioni medio-elevate, $d = 10$, valutando le performance del modello direttamente nel quantile 0.9, dove si erano evidenziate le maggiori criticità. Il processo è stato generato da un modello VAR(1). Per valutare se avesse senso o meno la stima di un modello in dimensione $d = 10$ è stato testato l'algoritmo 4.2.1. Sono state utilizzate 5 catene, con *burn-in* di 20000 osservazioni, estraendo una osservazione ogni 10 dalle 20000 osservazioni successive. Per il parametro \mathbf{D} si ha avuto una *effective sample size* di 337 osservazioni su un campione complessivo di 10000 elementi ed un *potential scale reduction factors* multivariato di 1.14. La stima del test LRcc è risultata la seguente:

algoritmo	quantile	LRcc	p-value
2	0.9	14.61	0.000

Viene riportato anche il grafico delle autocorrelazioni dei valori generati in figura 4.3. La scarsa qualità dei valori generati per \mathbf{D} nel quantile 0.9 influenza la capacità di stima del modello stesso; i valori di \mathbf{D} , entrando nel modello nella componente $\lambda \mathbf{v} \mathbf{D}$, che trasla la distribuzione nel quantile, se non adeguatamente stimati non permettono una corretta stima del quantile. La stima di un modello VAR, con dimensioni prossime a $d = 10$ o superiori, risulta quindi sconsigliabile, a meno di utilizzare catene estremamente lunghe e di usare per la stima meno di 1 valore ogni 10 generati.

4.3.3 Osservazioni finali

Lo studio di simulazione ha portato alla luce alcune criticità nella stima dei modelli VAR quantilici proposti. L'algoritmo 4.2.2, generando interamente la matrice $\mathbf{\Omega}$ tramite *Metropolis Hastings*, al crescere di d perde rapidamente qualità nella generazione delle catene; rimane tuttavia una alternativa valida per catene con $d \leq 4$. Una considerazione di carattere computazionale riguarda la stima di \mathbf{v} ; dovendo salvare in memoria n elementi ad ogni iterazione, la gestione dello spazio potrebbe diventare un fattore importante con n e numero di iterazioni maggiori. L'algoritmo *spike-and-slab* si è dimostrato in grado di fornire informazioni utili per valutare l'importanza dei coefficienti β , senza peggiorare le performance complessive del modello. Il modello con restrizioni

si è dimostrato come prevedibile una alternativa preferibile quando si hanno informazioni a priori sui coefficienti da stimare, situazione tuttavia rara nella pratica. La stima di \mathbf{D} effettuata tramite *Metropolis Hastings* diventa più complicata al crescere di d , influenzando pesantemente le performance dei modelli stimati, come visto aumentando la dimensione a $d = 6, 10$. I valori generati dalle catene per \mathbf{D} e $\boldsymbol{\mu}$ diventano più correlati al crescere di d , rendendo necessarie più osservazioni e la selezione di elementi dal campione a distanza maggiore l'uno dall'altro. L'implementazione di versioni più efficienti degli algoritmi, in particolare per quanto riguarda la generazione di \mathbf{D} , potrebbe permettere in futuro di applicare questi modelli anche nel caso di variabili risposta ad elevata dimensionalità, senza comprometterne la qualità delle stime.

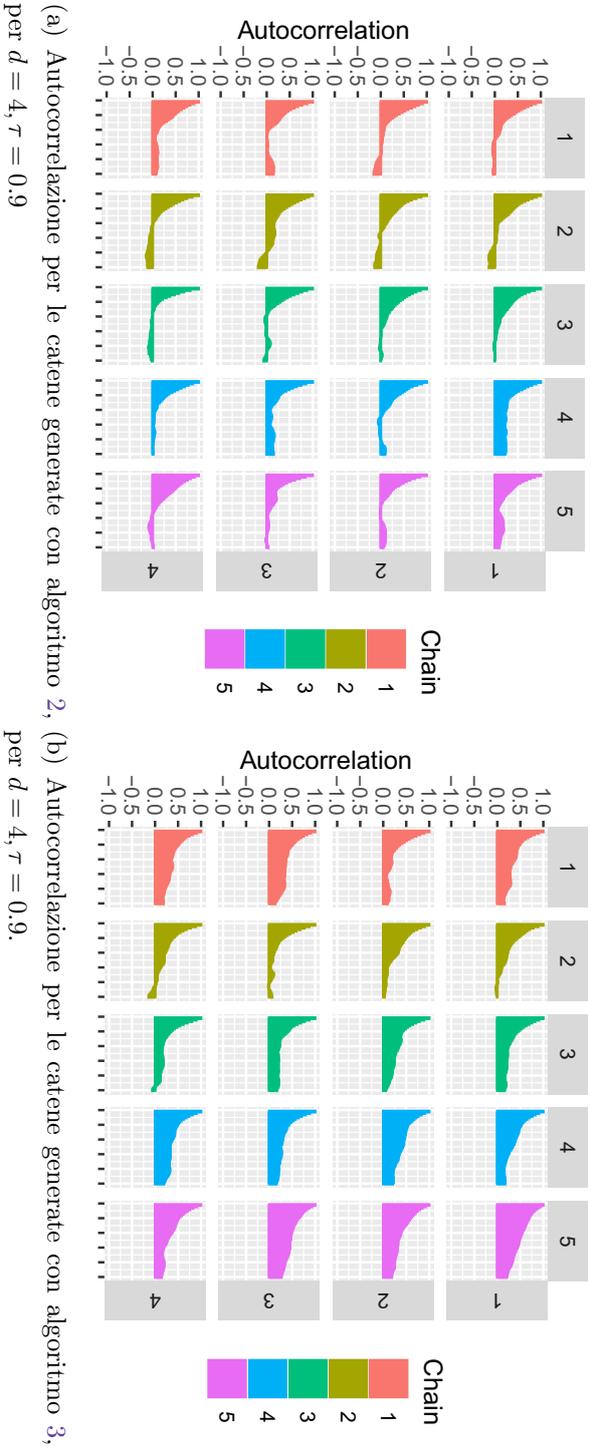
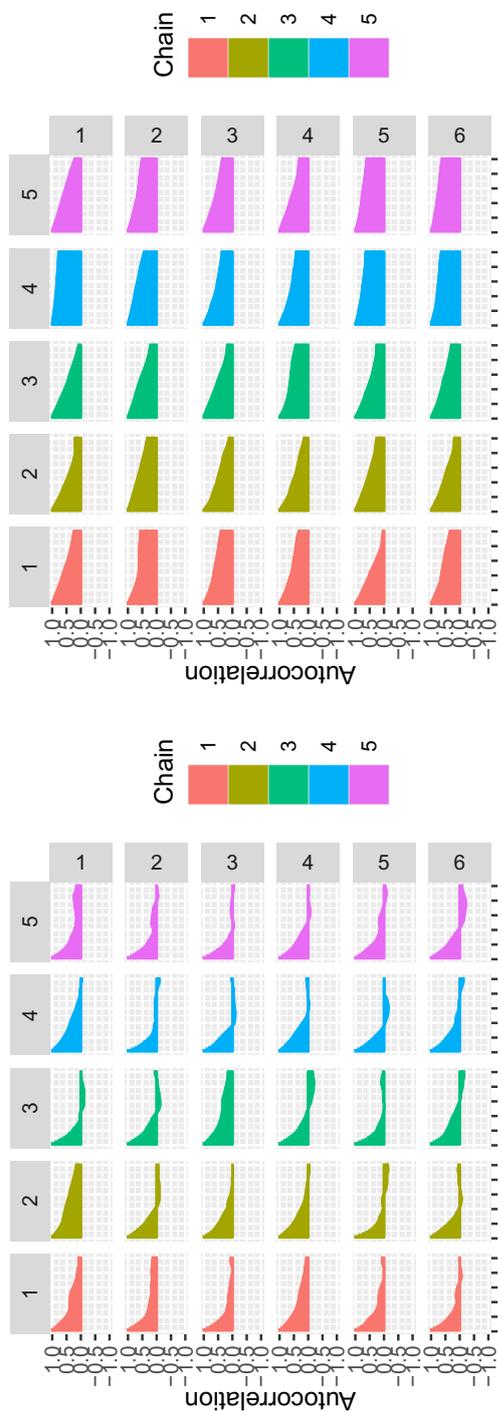


Figura 4.1: Confronto autocorrelazioni valori generati dalle catene MCMC per \mathbf{D} , $d = 4$.



(a) Autocorrelazione per le catene generate con algoritmo 2, (b) Autocorrelazione per le catene generate con algoritmo 3, per $d = 6, \tau = 0.9$.

Figura 4.2: Confronto autocorrelazioni valori generati dalle catene MCMC per $\mathbf{D}, d = 6$.

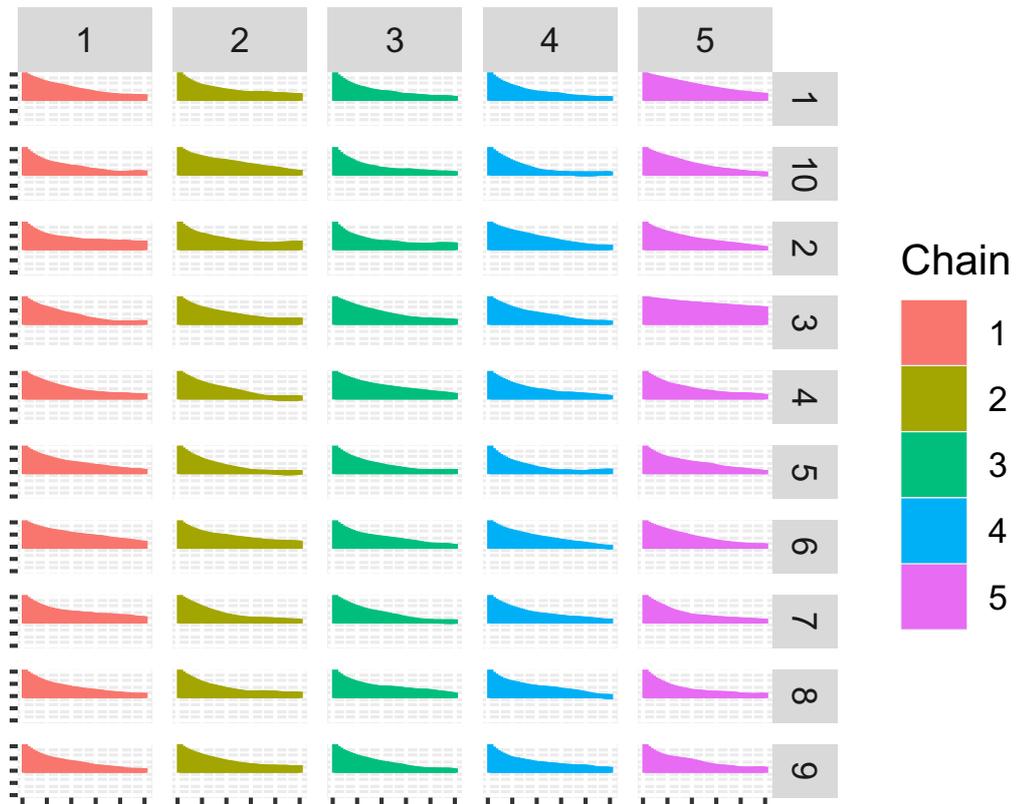


Figura 4.3: Autocorrelazione per le serie dei valori generati dalle catene MCMC per \mathbf{D} , dimensione $d = 10$.

Capitolo 5

Applicazione su dati reali

In questo capitolo vengono analizzati i prezzi dei *Credit Default Swap (CDS)* per un sottinsieme di paesi europei. Si vuole utilizzare i metodi e modelli presentati nel capitolo 3 per cercare di modellare ed evidenziare le differenti relazioni e dinamiche dei prezzi dei *CDS* delle principali economie dell'area Euro e calcolare gli effetti che uno shock su di un elemento avrebbe sull'evoluzione del sistema stesso. All'interno dell'Appendice A sono riportati alcuni grafici e misure sintetiche relative alle catene generate dagli algoritmi stimati in questo capitolo.

5.1 Misure di rischio

Da fine anni '80 in poi vi è stata una crescente attenzione da parte delle istituzioni alla stabilità e solidità degli istituti finanziari, attenzione che ha portato alla stesura di accordi internazionali in tema di valutazione del rischio per gli istituti finanziari tra i paesi del G10, Basilea(1988), poi aggiornati nel 2004 e nel 2010 con Basilea II e Basilea III. Uno dei principali strumenti utilizzati per la stima del rischio da parte degli istituti finanziari è il *value at risk*. Dato un *asset*, o un insieme di *asset*, all'interno di un portafoglio, il *value at risk* a livello τ , è definito come la perdita massima stimata in un orizzonte temporale h fissato, ad un livello di confidenza τ . Definiamo quindi

il *value at risk* a livello τ , per un *asset* \mathbf{Y} al tempo i , con orizzonte h , come:

$$P(\mathbf{Y}_{i+h} \leq VaR_{i+h,\tau}) = \tau.$$

Per poter calcolare il *value at risk* è quindi necessario stimare la distribuzione dell'*asset* e poterne calcolare il quantile τ . I modelli quantilici risultano immediatamente applicabili al calcolo del *value at risk*, in quanto permettono la stima della distribuzione dell'*asset* direttamente nel quantile τ di interesse. La modellazione quantilica multivariata, come il modello VAR quantilico introdotto, offre la possibilità non solo di stimare la dinamica del *value at risk* relativo ad un solo asset, ma di calcolare le relazioni che intercorrono tra più asset di interesse. Per valutare il *value at risk* di un asset \mathbf{Y}^l , condizionatamente al valore di un asset \mathbf{Y}^j , sono state proposte vari metodi, tra cui il *conditional value at risk*, definito come il *value at risk* a livello τ , con orizzonte h , per una variabile l , condizionatamente ad un determinato evento sulla variabile j :

$$P(\mathbf{Y}_{i+h}^l | \mathbf{Y}_i^j \leq CoVaR_{jl,i+h,\tau}) = \tau.$$

Il concetto di *conditional value at risk* è strettamente legato al modello VAR quantilico introdotto. Un modello VARX quantilico, come quello introdotto in questa tesi, può essere interpretato come il calcolo del *conditional value at risk* a livello τ , con $h = 1$, orizzonte temporale, per le variabili modellate in \mathbf{y}_i . Dato il modello VARX(p) introdotto nel capitolo 3:

$$\mathbf{y}_i = \boldsymbol{\mu} + \mathbf{A}_1 \mathbf{y}_{i-1} + \cdots + \mathbf{A}_p \mathbf{y}_{i-p} + \mathbf{A}_x \mathbf{x}_i + \boldsymbol{\varepsilon}_i,$$

con $\boldsymbol{\varepsilon}_i \sim AL_d(0, \mathbf{DRD}, \tau)$, termine di errore in grado di centrare il modello in τ , si può riscrivere il modello in termini di probabilità come:

$$P(\mathbf{y}_{i+1} | (\mathbf{y}_i, \dots, \mathbf{y}_{i-p}, \mathbf{x}_i) \leq \boldsymbol{\mu} + \mathbf{A}_1 \mathbf{y}_{i-1} + \cdots + \mathbf{A}_p \mathbf{y}_{i-p} + \mathbf{A}_x \mathbf{x}_i) = \tau,$$

con il termine $\boldsymbol{\mu} + \mathbf{A}_1 \mathbf{y}_{i-1} + \cdots + \mathbf{A}_p \mathbf{y}_{i-p} + \mathbf{A}_x \mathbf{x}_i$ che può essere interpretato come il *conditional value at risk* per \mathbf{y}_1 a ritardo $h = 1$. La stima del *value at risk* e varianti è solo uno degli strumenti utilizzati per la valutazione del rischio. Spesso risulta di interesse valutare le performance di un *asset* in una specifica situazione di crisi; per fare questo negli anni sono stati introdotti diversi tipi di *stress test*. Con *stress test* si indica generalmente un insieme di

simulazioni volte a valutare l'esposizione al rischio di un portafoglio di *asset* o di un istituto, al verificarsi di determinate condizioni. La regressione quantilica, con particolare attenzione alla modellazione VAR quantilica, può tornare utile nella simulazione di uno *stress test*, grazie alla possibilità di modellare il comportamento del portafoglio direttamente nel quantile τ associato alla situazione di stress, usando quest'informazione per simulare una situazione di stress relativa ad osservazioni sul quantile τ . In questo capitolo verrà ripreso il metodo di Chavleishvili e Manganeli 2019 per la creazione di uno *stress test*, generando osservazioni successive dal processo in un quantile τ associato ad una situazione di stress, seguito poi da osservazioni generate in mediana, così da valutare, in seguito alla situazione di stress, i cambiamenti avvenuti al sistema analizzato.

5.2 Dataset *CDS*

I *Credit Default Swap* sui titoli di stato sono un derivato finanziario, in cui chi acquista il *CDS*, a fronte di una serie di pagamenti, ha diritto ad essere rimborsato del valore del sottostante, titoli di stato, qualora questi non risultino più esigibili. I *CDS* possono essere considerati come una assicurazione sul rischio di *default* di un paese. Il prezzo dei *CDS* riflette la percezione del rischio sul debito sovrano di un paese; il prezzo dei *CDS* ci si aspetta aumenti all'aumentare del rischio di *default* di un paese e che diminuisca al diminuire del rischio; di conseguenza il prezzo dei *CDS* può essere una valida misura del rischio paese percepito dagli investitori. Per approfondimenti sulle caratteristiche dei *CDS* e sulle loro dinamiche fare riferimento a Naifar 2020. Il dataset di riferimento, ottenuto tramite Eikon., contiene i prezzi dei *CDS* per i paesi dell'Unione Europea e per il resto delle principali economie mondiali e ha cadenza settimanale; il dataset copre il periodo dal "25/12/2007" fino al "28/07/2020". Il dataset di riferimento contiene le osservazioni relative ai prezzi dei *CDS* di 57 paesi. Nel seguito si focalizzerà l'analisi sulle serie dei prezzi dei *CDS* relativi ad alcune economie europee selezionate, con l'eventuale inserimento di covariate esterne nel modello.

Le serie dei prezzi risultano difficilmente trattabili direttamente, per via della presenza di coefficienti molto vicini a radici unitarie, che rendono le serie non

stazionarie e difficilmente modellabili con modelli VAR. La soluzione tipicamente adottata in questi casi è quella di trattare la serie dei log-rendimenti invece che dei prezzi. I log-rendimenti percentuali (r) sono definiti in funzione dei prezzi (p) dei *CDS* come $r_i = \log\left(\frac{p_i}{p_{i-1}}\right)100$. Viene riportato un grafico di esempio contenente le funzioni di autocorrelazione e autocorrelazione parziale stimate, per la serie dei prezzi dei *CDS* relativi all'Italia, e la serie dei log-rendimenti percentuali per l'Italia. Si può notare nei grafici, figure 5.1, 5.2, che la trasformazione in log-rendimenti risolve il problema della persistenza della autocorrelazione nella serie. Il dataset, coprendo il periodo "25/12/2007"-

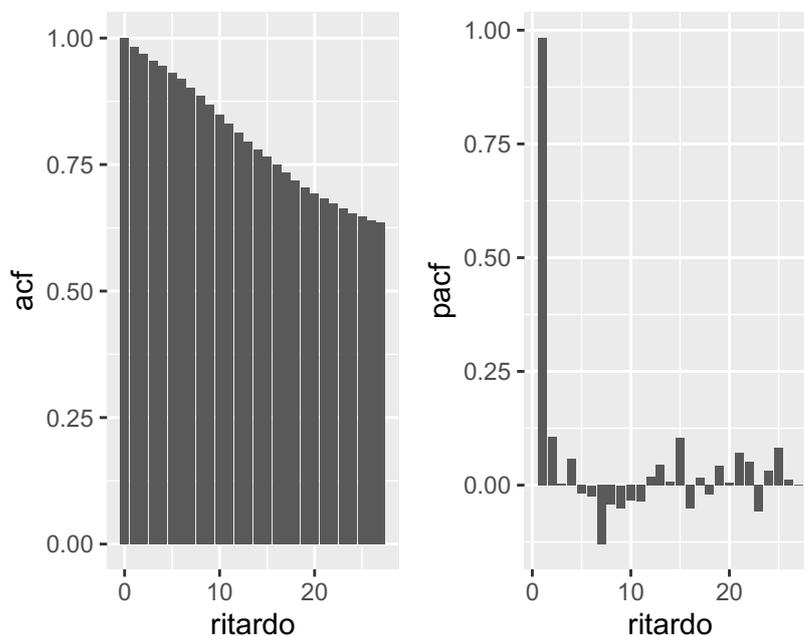


Figura 5.1: Funzione di autocorrelazione e di autocorrelazione parziale per le serie dei prezzi dei *CDS* per l'Italia.

"28/07/2020", contiene al suo interno anche i dati relativi alla cosiddetta "Crisi della zona euro", periodo 2011-2014 indicativamente. Date le implicazioni che tale crisi ha avuto sul sistema Euro, oltre che l'entrata in vigore di politiche di *Quantitative Easing* da parte della Banca Centrale Europea, con il relativo impatto sul prezzo dei *CDS*, e piani di aiuti economici per la Grecia, risulta difficile modellare l'intero periodo "25/12/2007"- "28/07/2020" come generato dallo stesso processo. Di conseguenza l'analisi è stata divisa in

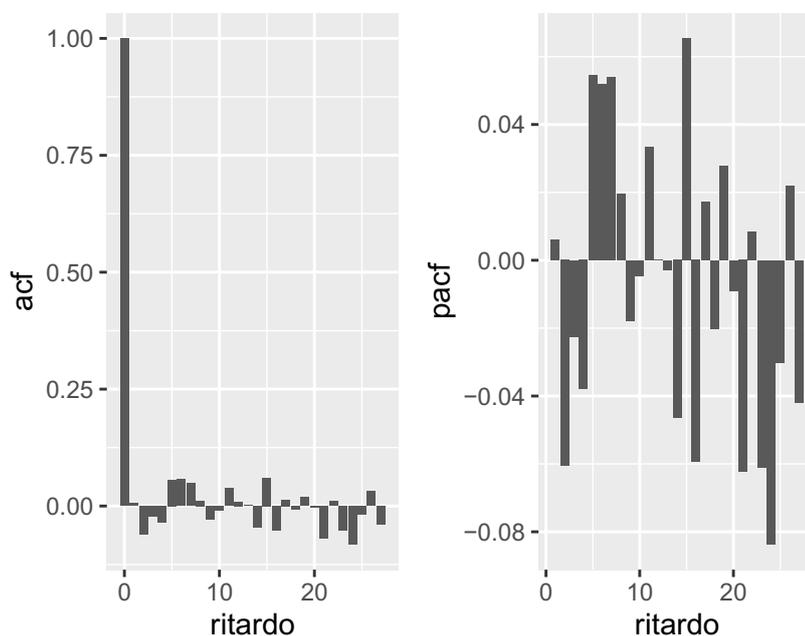


Figura 5.2: Funzione di autocorrelazione e di autocorrelazione parziale per le serie dei log-rendimenti percentuali dei *CDS* per l'Italia.

due parti; la prima analisi riguarda il sistema dei *CDS* pre-crisi, modellando le serie relative a Germania, Italia e Grecia, la seconda analisi riguarda il sistema post-crisi; vengono modellate le serie relative a Germania, Francia, Spagna, Italia e Grecia, usando come covariate le serie relative a Stati Uniti, Belgio, Olanda, Portogallo.

5.3 Analisi Italia, Grecia, Germania

Una prima analisi affronta la modellazione delle serie dei log-rendimenti relativi ai *CDS* di Italia, Grecia e Germania. In questa analisi vengono modellati i dati relativi al periodo che va dal “25/12/2007” al “20/04/2010”. Non è stato possibile reperire una serie di prezzi dei *CDS* relativi ai debiti sovrani precedente al “25/12/2007”; ho quindi scelto di modellare congiuntamente solamente Italia, Grecia e Germania in questa prima parte per via del numero ridotto di osservazioni a disposizione. La data del “20/04/2010” è stata scelta per tenere conto delle tempistiche storiche relative alla crisi del debi-

to Greco, con il primo dei *downgrade* dei titoli di stato greci da parte delle principali agenzie di *rating* che sarebbe arrivato nel corso della settimana successiva. L'obiettivo di questa analisi è quello di valutare, utilizzando la modellazione VAR quantilica introdotta, se il comportamento delle serie nel periodo successivo risulti essere coerente con le dinamiche precedenti e se, tramite opportuni strumenti, l'evoluzione del sistema potesse essere in parte predetta con i dati a disposizione. La serie è stata modellata con un modello VAR(2) nei quantili (0.1, 0.5, 0.9). La stima del modello è avvenuta tramite algoritmo MCMC, per ogni quantile, con 10000 iterazioni di *burn-in*, estraendo 1 elemento su 10 dalle successive 10000 iterazioni come campione. Per ogni quantile sono state stimate 3 catene in parallelo, combinandone poi i valori. Dato il numero di parametri viene qui riportato solamente il *potential scale reduction factors* multivariato (PSRF) come misura riassuntiva per i parametri (β, Ω, D, ν) ; in Appendice A vengono riportate alcune misure ulteriori relative alle catene MCMC da cui sono stati generati i 3 modelli.

quantile	β	Ω	D	ν
0.1	1.08	1.08	1.05	1.09
0.5	1.01	1.06	1.05	1.07
0.9	1.05	1.02	1.07	1.09

Tabella 5.1: *Potential scale reduction factors* multivariato per le catene MCMC generate.

Di seguito vengono riportati i 3 modelli VAR stimati, nei 3 quantili analizzati. Il primo modello stimato, relativo al quantile $\tau = (0.1, 0.1, 0.1)$, risulta:

$$\begin{bmatrix} Grecia_i \\ Italia_i \\ Germania_i \end{bmatrix} = \begin{bmatrix} -12.51 \\ -13.37 \\ -16.28 \end{bmatrix} + \begin{bmatrix} 0.33 & 0.46 & 0.32 \\ -0.20 & -0.33 & -0.24 \\ -0.00 & -0.02 & 0.05 \end{bmatrix} \begin{bmatrix} Grecia_{i-1} \\ Italia_{i-1} \\ Germania_{i-1} \end{bmatrix} + \begin{bmatrix} 0.57 & 0.59 & 0.71 \\ -0.48 & -0.62 & -0.72 \\ 0.05 & 0.13 & 0.09 \end{bmatrix} \begin{bmatrix} Grecia_{i-2} \\ Italia_{i-2} \\ Germania_{i-2} \end{bmatrix}.$$

Per $\tau = (0.5, 0.5, 0.5)$ il modello stimato risulta:

$$\begin{bmatrix} Grecia_i \\ Italia_i \\ Germania_i \end{bmatrix} = \begin{bmatrix} 0.74 \\ -0.04 \\ 0.01 \end{bmatrix} + \begin{bmatrix} 0.26 & 0.43 & 0.24 \\ -0.16 & -0.31 & -0.17 \\ -0.03 & -0.05 & -0.02 \end{bmatrix} \begin{bmatrix} Grecia_{i-1} \\ Italia_{i-1} \\ Germania_{i-1} \end{bmatrix} + \begin{bmatrix} 0.25 & 0.32 & 0.37 \\ -0.24 & -0.39 & -0.45 \\ -0.01 & 0.05 & 0.03 \end{bmatrix} \begin{bmatrix} Grecia_{i-2} \\ Italia_{i-2} \\ Germania_{i-2} \end{bmatrix}.$$

Per $\tau = (0.9, 0.9, 0.9)$ il modello stimato risulta:

$$\begin{bmatrix} Grecia_i \\ Italia_i \\ Germania_i \end{bmatrix} = \begin{bmatrix} 16.84 \\ 16.08 \\ 18.59 \end{bmatrix} + \begin{bmatrix} 0.43 & 0.58 & 0.44 \\ -0.19 & -0.34 & -0.30 \\ -0.16 & -0.17 & -0.12 \end{bmatrix} \begin{bmatrix} Grecia_{i-1} \\ Italia_{i-1} \\ Germania_{i-1} \end{bmatrix} + \begin{bmatrix} 0.10 & 0.17 & 0.20 \\ -0.30 & -0.48 & -0.49 \\ 0.09 & 0.17 & 0.11 \end{bmatrix} \begin{bmatrix} Grecia_{i-2} \\ Italia_{i-2} \\ Germania_{i-2} \end{bmatrix}.$$

Confrontando i *drift* dei 3 paesi nei 3 quantili, la Germania sembra avere una variabilità media maggiore al netto dell'effetto dei ritardi; è importante ricordare che si tratta di variazioni sui log-rendimenti, non sui prezzi stessi, tendenzialmente più stabili per la Germania. Confrontando i coefficienti dei 3 modelli sembra esserci una coerenza, nei 3 quantili di riferimento, dell'effetto dei ritardi. Si può notare dai grafici in figura 5.3 come i log-rendimenti abbiano un periodo di forti variazioni positive nel Settembre-Ottobre 2008, periodo coincidente con il fallimento della banca americana Lehman Brothers.

Vengono riportate le matrici di covarianza delle innovazioni stimate per i 3 quantili:

$\tau = 0.9$	<i>Gre</i>	<i>Ita</i>	<i>Ger</i>	$\tau = 0.5$	<i>Gre</i>	<i>Ita</i>	<i>Ger</i>	$\tau = 0.1$	<i>Gre</i>	<i>Ita</i>	<i>Ger</i>
<i>Gre</i>	62.47	54.07	57.51	<i>Gre</i>	42.74	31.37	32.66	<i>Gre</i>	43.06	33.54	38.07
<i>Ita</i>	54.07	62.50	60.27	<i>Ita</i>	31.37	39.42	35.83	<i>Ita</i>	33.54	41.66	40.82
<i>Ger</i>	57.51	60.27	82.69	<i>Ger</i>	32.66	35.83	62.78	<i>Ger</i>	38.07	40.82	65.72

Tabella 5.2: Matrici di covarianza stimate, generate tramite algoritmo MCMC.

le 3 matrici di covarianza stimate hanno correlazioni simili nei 3 periodi, con una varianza maggiore tuttavia nel quantile 0.9. Per poter valutare se il comportamento della serie, nei quantili stimati, è rimasto invariato nell'anno e mezzo successivo, periodo che intercorre tra l'inizio della crisi Greca ed il momento di massima crisi, con l'esplosione dei rendimenti dei titoli di stato greci ed ulteriori *downgrade* da parte delle agenzie di *rating*, sono state effettuate delle previsioni ad una settimana con il modello VAR(2) stimato. Valutare se il comportamento non subisce variazioni significative equivale ad assumere che la stima del *conditional value at risk* effettuata dal modello sia valida ed utilizzabile anche nel periodo successivo. Per valutare l'adeguatezza dei quantili stimati è stato utilizzato il test LRcc Christoffersen 1998.

quantile	Grecia	Italia	Germania	LRcc	p-value
0.1	0.102	0.170	0.0681	0.644	0.724
0.5	0.465	0.511	0.568	2.962	0.227
0.9	0.898	0.886	0.932	1.425	0.490

I test non rifiutano l'ipotesi nulla che il *value at risk* previsto, stimato tramite modello per la settimana successiva, coincida con il quantile equivalente

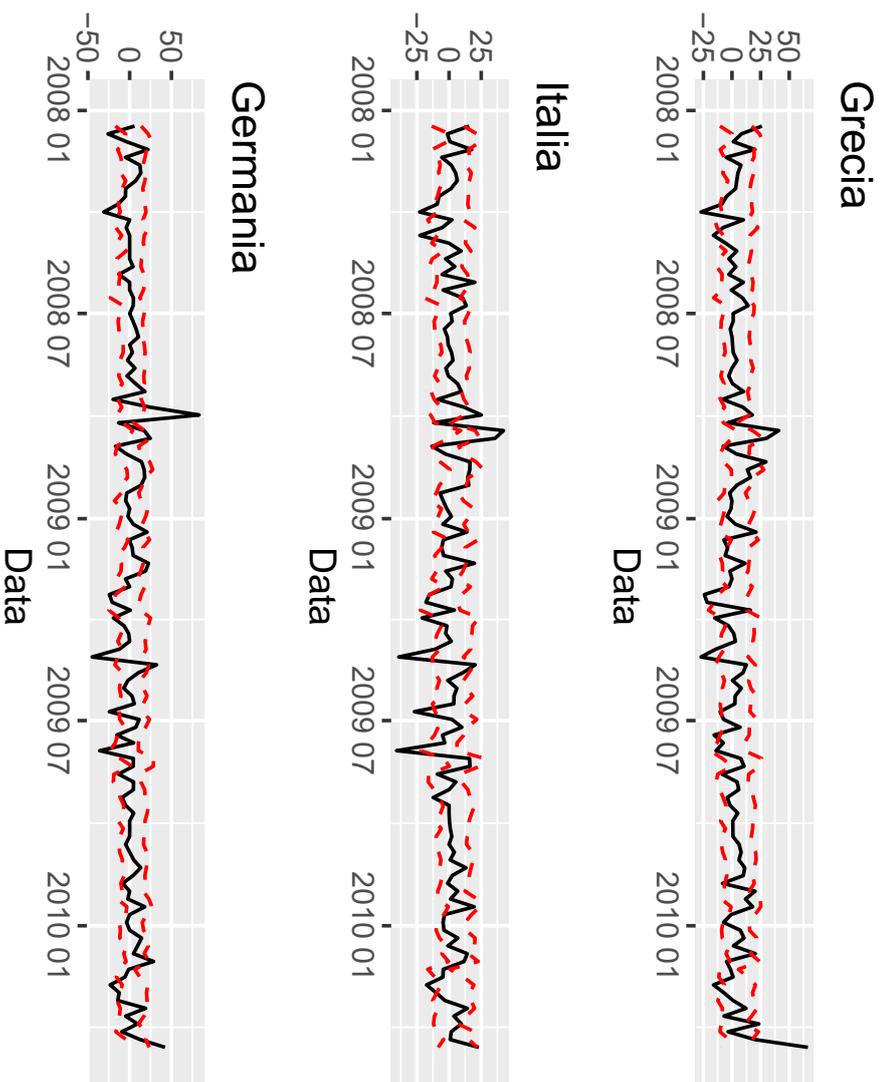


Figura 5.3: Serie osservata dei log-rendimenti dei *CDS* nel periodo di stima, con la relativa stima dei quantili (0.1,0.9) tramite i modelli VAR quantilici.

dei dati. Si può quindi affermare che il *value at risk* stimato si adatta bene ai dati per $\tau = 0.1, 0.5, 0.9$, e che il modello riesca a prevedere in maniera adeguata il comportamento della serie nei quantili studiati. Di seguito vengono riportati, come termine di paragone, i risultati ottenuti nel medesimo periodo, utilizzando un *value at risk* costante, definito sui quantili stimati nel periodo precedente.

quantile	Grecia	Italia	Germania	LRcc	p-value
0.1	0.102	0.125	0.0455	0.581	0.748
0.5	0.443	0.534	0.511	0.018	0.990
0.9	0.909	0.897	0.954	1.380	0.501

Confrontandolo con le performance ottenute tramite *value at risk* stimato staticamente il modello utilizzato non mostra importanti differenze o vantaggi. Entrambi i *value at risk* stimati suggeriscono un lieve spostamento verso l'alto per i rendimenti mediani greci, ed una minore variabilità dei rendimenti tedeschi, con una proporzione di elementi inferiori al quantile 0.1 e superiore al quantile 0.9 minore delle attese. Viene riportato in figura 5.4 il grafico delle serie relativo al periodo Maggio 2010-Dicembre 2011, con i quantili 0.1, 0.9 stimati tramite modello VAR.

5.3.1 Analisi strutturali quantiliche

Per poter analizzare meglio le dinamiche della serie sono stati utilizzati strumenti di analisi strutturale quantilica, come le *generalised quantile forecast error variance decomposition* e *quantile impulse response function* definite in 3.5.1. Tramite *generalised quantile forecast error variance decomposition* è possibile stimare, ad un ritardo fissato, come gli *shock* su ciascuna delle tre variabili concorrano alla varianza dell'errore di previsione per una variabile, dandoci così una idea di quali siano le variabili a guidare maggiormente le variazioni del sistema. Per l'analisi tramite *generalised quantile forecast error variance decomposition* è necessario scegliere a che ritardo operare la stima; trattandosi di serie settimanali si è scelto di stimare le QGFEVD a 4 settimane. Vengono riportati gli intervalli di fiducia di probabilità 0.8, per ogni elemento delle QGFEVD stimate; l'intervallo è stato ottenuto calcolando le

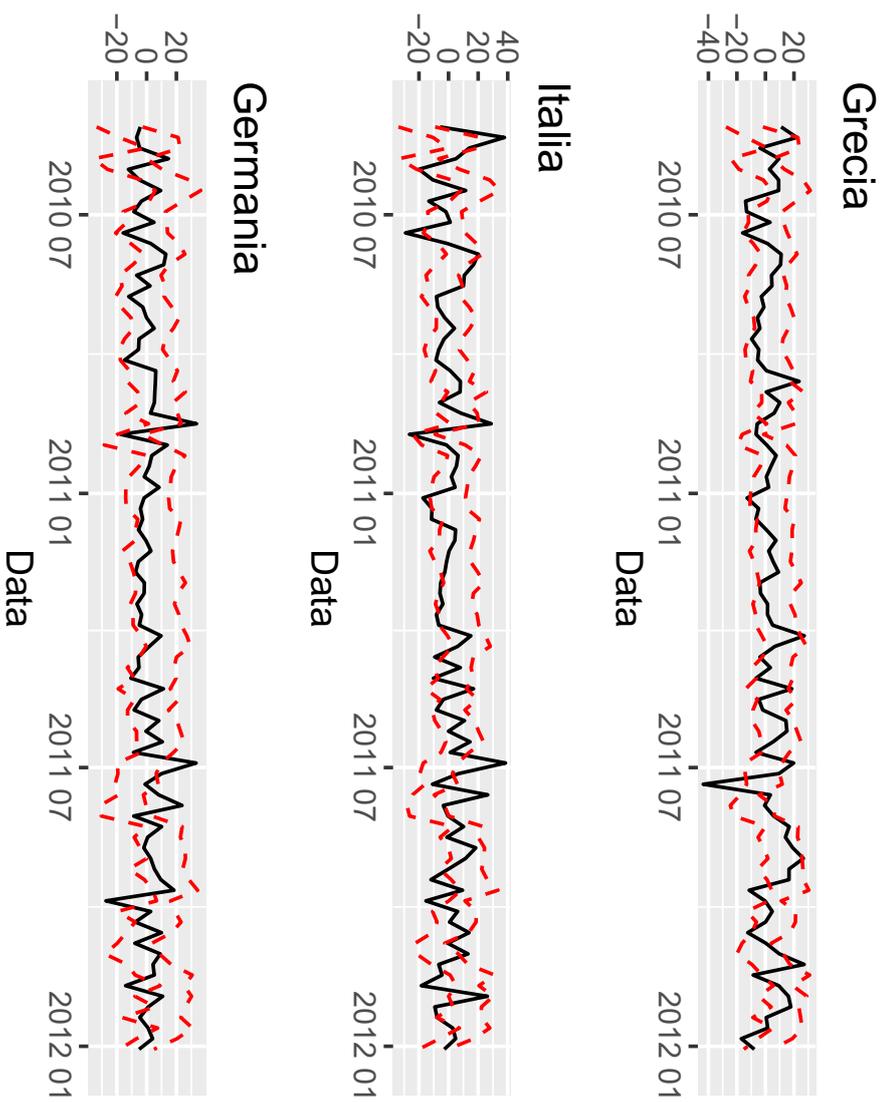


Figura 5.4: Serie osservata dei log-rendimenti dei *CDS* nel periodo Maggio 2010-Dicembre 2011, con la relativa stima dei quantili (0.1,0.9) ottenute tramite i modelli VAR quantilici stimati.

QGFEVD per ogni estrazione dalla catena, ed estraendo i quantili empirici 0.1 e 0.9. Le QGFEVD qui riportate, tabelle 5.3, 5.4, 5.5, vanno lette per colonna, ed indicano una stima di quanto la variabilità degli errori di previsione sulla variabile di colonna dipenda dalle innovazioni sulle variabili di riga. Le

quantile 0.1	Grecia	Italia	Germania
Grecia	0.46-0.71	0.39-0.72	0.30-0.76
Italia	0.22-0.42	0.17-0.44	0.13-0.43
Germania	0.02-0.15	0.02-0.23	0.02-0.33

Tabella 5.3: *generalised quantile forecast error variance decomposition*, 4 settimane, quantile 0.1.

quantile 0.5	Grecia	Italia	Germania
Grecia	0.46-0.72	0.37-0.73	0.26-0.77
Italia	0.23-0.44	0.19-0.49	0.11-0.48
Germania	0.01-0.11	0.01-0.2	0.01-0.33

Tabella 5.4: *Generalised quantile forecast error variance decomposition*, 4 settimane, quantile 0.5.

quantile 0.9	Grecia	Italia	Germania
Grecia	0.47-0.71	0.37-0.74	0.44-0.76
Italia	0.23-0.42	0.16-0.46	0.16-0.32
Germania	0.02-0.13	0.02-0.22	0.02-0.21

Tabella 5.5: *Generalised quantile forecast error variance decomposition*, 4 settimane, quantile 0.9.

generalised quantile forecast error variance decomposition, suggeriscono l'idea che la variabile Grecia sia quella a guidare maggiormente le innovazioni nel processo, sia per se stessa che per Italia e Germania. Per valutare invece come il sistema si evolve, in caso di un singolo shock su una variabile, è possibile utilizzare la *quantile impulse response function*. Vengono utilizzate le *quantile impulse response function* definite in 3.5.1 per valutare l'effetto che

uno *shock* sui prezzi dei *CDS* greci ha sullo sviluppo del sistema nelle settimane successive. Per ogni nazione e per ogni quantile $\tau = (0.1, 0.5, 0.9)$ viene riportata la relativa *quantile impulse response function*, con i relativi intervalli di fiducia di livello 0.8, figura 5.5. Le *quantile impulse response function* suggeriscono una maggiore velocità di assorbimento degli *shock* per i *CDS* tedeschi, con i prezzi dei *CDS* italiani che sembrano invece avere un aumento di variabilità nelle settimane successive, con oscillazioni dei rendimenti maggiori; per i *CDS* greci al contrario sembra che uno *shock* comporti aumenti dei log-rendimenti anche nelle settimane successive, e che venga più difficilmente riassorbito. Le *quantile impulse response function* possono aiutarci a capire la reazione del sistema ad un singolo *shock*, tuttavia può risultare di interesse valutare come il sistema reagirebbe se posto sotto maggiore tensione, con una serie di *shock* successivi. Similmente a quanto fatto da Chavleishvili e Manganelli 2019, è stato effettuato uno *stress test*, creando una situazione di *stress* sul rendimento dei *CDS* greci, e valutando come questo modificasse il sistema. Per fare *stress test* sono stati simulati i log-rendimenti greci per 6 settimane sul quantile 0.9, generando invece Germania e Italia nel quantile 0.5; nelle 6 settimane successive ho riportato il sistema in una situazione normale, generando tutte e 3 le serie nel quantile 0.5 e valutando come era cambiato il sistema. Le serie dei prezzi sono state ottenute applicando i log-rendimenti simulati all'ultimo prezzo a disposizione nel periodo di stima. Lo *stress test*, riportato in figura 5.6, prevede un aumento dei prezzi molto elevato per i *CDS* greci, con un aumento maggiore per i *CDS* Italiani rispetto a quelli tedeschi. È interessante notare come i livelli dei prezzi raggiunti dallo *stress test* siano molto vicini ai veri valori che si sarebbero visti nell'anno successivo, a metà 2011, in piena crisi Greca; per esempio i prezzi dei *CDS* al "26/7/2011" risultavano pari a 1826 per la Grecia, 272 per l'Italia, 59 per la Germania, valori simili a quelli visti nello *stress test* nel caso di stress sulla variabile greca. Si può quindi dire che l'esplosione dei prezzi avvenuta successivamente al peggioramento della situazione greca fosse già insita nel processo, e si potesse valutare tramite un'opportuna modellazione delle serie a disposizione ed una opportuna analisi.

5.4 Analisi CDS post-crisi

In questa sezione viene svolta una seconda analisi sui prezzi dei CDS relativi a paesi europei, focalizzandosi questa volta su un periodo più recente. Dopo aver visto nella precedente sezione come i modelli quantilici introdotti sembrano stimare adeguatamente il comportamento della serie dei log-rendimenti, sia nel periodo di stima che nel periodo di crisi successiva, può risultare di interesse utilizzare questi strumenti per valutare le caratteristiche e le dinamiche dei prezzi dei CDS per i titoli di stato di alcuni paesi europei selezionati in un periodo più recente, post "Crisi della zona euro". In questa sezione vengono modellate le serie dei log-rendimenti dei CDS relative al periodo "16/12/2014"-“24/12/2019”, studiando e modellando congiuntamente il comportamento delle serie relative a: Italia, Francia, Spagna, Grecia, Germania. Sono state utilizzate come covariate le serie dei log-rendimenti, a ritardo 1, di Stati Uniti, Portogallo, Belgio e Olanda. La serie dei log-rendimenti per gli Stati Uniti è stata inserita così da avere un termine di confronto esterno all'area Euro, scegliendo appunto i CDS degli Stati Uniti per l'influenza che hanno su scala globale. Anche in questo caso è stato scelto di modellare la serie nei quantili $\tau = (0.1, 0.5, 0.9)$

5.4.1 Modelli stimati

Come già accennato nei precedenti capitoli, la scelta del ritardo p , oltre che delle covariate da inserire, è una scelta importante per la stima del modello. A differenza di quanto avviene per la modellazione VAR in media, non esiste un metodo immediato per la stima del ritardo all'interno di un modello VAR(p) quantilico bayesiano. Per selezionare un modello VAR che si adattasse il meglio possibile ai dati ho deciso di implementare e confrontare tra loro 3 possibili modelli tramite una procedura di *backtesting*, valutando la capacità dei modelli di stimare correttamente i quantili nell'insieme di stima. Il primo modello proposto è un VARX(2) quantilico, contenente quindi due ritardi per i log-rendimenti modellati, oltre alle covariate. Il secondo modello proposto è un modello VARX(2) quantilico con priori di tipo *spike-and-slab*, con i seguenti iperparametri ($a_\pi = 15, b_\pi = 5, k_0 = 0.0001, a_k = 1, b_k =$

1, $a_\sigma = 1, b_\sigma = 1$). Il terzo modello proposto è un modello VARX(3) con restrizioni, in cui si è ristretto \mathbf{A}_2 e \mathbf{A}_3 , matrici contenenti i coefficienti di regressioni per i termini autoregressivi a ritardo 2 e 3, ad una forma diagonale. Come termine di paragone per valutare le performance dei 3 modelli è stato utilizzato un modello lineare quantilico sulle singole componenti della risposta, equivalente allo stimare una alla volta le righe del modello VAR(2) quantilico proposto. Le performance dei modelli vengono valutate utilizzando il test LRcc, per valutare l'adeguatezza nella stima del quantile all'interno dell'insieme di stima. Il modello VAR(3) con restrizioni risulta poco per-

modello	quantile	quantile stimato	LRcc	p-value
VARX(2)	0.1	0.087	3.532	0.171
VARX(2)	0.5	0.508	6.708	0.034
VARX(2)	0.9	0.912	1.992	0.158
VARX(3) con restrizioni	0.1	0.122	8.367	0.015
VARX(3) con restrizioni	0.5	0.504	7.581	0.023
VARX(3) con restrizioni	0.9	0.926	11.784	0.003
VARX(2) <i>Spike and slab</i>	0.1	0.108	2.070	0.355
VARX(2) <i>Spike and slab</i>	0.5	0.509	5.954	0.051
VARX(2) <i>Spike and slab</i>	0.9	0.876	10.221	0.006
regressione lineare quantilica	0.1	0.118	9.577	0.008
regressione lineare quantilica	0.5	0.505	4.458	0.106
regressione lineare quantilica	0.9	0.889	3.706	0.156

Tabella 5.6: Tabella riassuntiva per la valutazione delle capacità di stima dei modelli proposti tramite test LRcc.

formante in tutti e tre i quantili. Al contrario il modello VAR(2) con priori *spike-and-slab* e il modello VAR(2) base mostrano un ottimo adattamento nei quantili 0.1 e 0.9 rispettivamente. Nel quantile 0.5 tutti e quattro i modelli utilizzati risultano ben centrati sul quantile; i valori elevati del test LRcc, definito come $LRcc = LRind + LRuc$, dipendono dalla componente LRind, che misura la dipendenza a ritardo 1 della copertura quantilica. Per modellare la serie è stato scelto di utilizzare un modello VARX(2) per il quantile 0.9 ed un modello VARX(2) con priori *spike-and-slab* per i quantili 0.5, 0.1. Di seguito vengono riportati i modelli stimati. Grafici e misure sintetiche sulle

catene generate dagli algoritmi di stima sono riportate in Appendice A.

Per il quantile $\tau = 0.9$ il modello stimato risulta:

	<i>drift</i>	Ger_{i-1}	Spa_{i-1}	Fra_{i-1}	Gre_{i-1}	Ita_{i-1}	Ger_{i-2}	Spa_{i-2}	Fra_{i-2}	Gre_{i-2}	Ita_{i-2}	Bel_{i-1}	Ola_{i-1}	Por_{i-1}	USA_{i-1}
Ger_i	7.867	-0.108	-0.045	0.050	-0.004	0.181	-0.050	0.009	0.090	0.028	0.025	0.028	-0.418	0.019	-0.041
Spa_i	9.965	-0.024	-0.025	0.012	-0.067	0.135	-0.005	0.077	0.046	-0.003	-0.119	0.081	-0.466	0.005	0.054
Fra_i	9.058	-0.028	-0.038	0.062	-0.034	0.186	0.003	0.102	0.033	0.015	-0.076	-0.028	-0.321	0.019	-0.088
Gre_i	5.042	0.005	0.037	-0.076	-0.047	0.199	0.046	0.053	0.013	0.032	-0.007	0.060	-0.230	0.004	0.023
Ita_i	10.280	-0.068	-0.028	0.023	-0.085	0.168	-0.032	0.099	-0.003	0.017	-0.093	0.051	-0.403	0.011	-0.040

Il modello stimato nel quantile $\tau = 0.5$ risulta:

	<i>drift</i>	Ger_{i-1}	Spa_{i-1}	Fra_{i-1}	Gre_{i-1}	Ita_{i-1}	Ger_{i-2}	Spa_{i-2}	Fra_{i-2}	Gre_{i-2}	Ita_{i-2}	Bel_{i-1}	Ola_{i-1}	Por_{i-1}	USA_{i-1}
Ger_i	0.103	-0.027	-0.012	0.060	0.006	0.025	-0.010	-0.008	0.053	0.001	0.012	0.002	-0.010	0.000	-0.006
Spa_i	-0.615	-0.001	0.013	0.011	-0.021	-0.019	0.003	-0.005	-0.002	-0.003	-0.058	-0.001	-0.031	-0.001	0.018
Fra_i	-0.178	0.030	0.032	0.039	-0.000	0.033	0.012	0.022	0.017	-0.001	-0.057	-0.046	0.028	0.003	0.003
Gre_i	-1.109	0.010	0.015	-0.012	-0.004	0.039	0.018	0.006	0.011	0.004	0.003	0.008	0.015	0.001	0.028
Ita_i	-0.433	-0.004	0.017	0.004	-0.036	-0.003	-0.015	-0.026	-0.027	0.006	-0.038	-0.010	0.001	-0.000	0.006

La stima a posteriori delle probabilità di appartenere alla componente *slab*, ricavata dalla stima a posteriori delle proporzioni di elementi $\gamma_{ij} = 1$, che può essere interpretata come una misura della significatività dell'apporto della variabile, risulta essere per $\tau = 0.5$:

	π	Ger_{i-1}	Spa_{i-1}	Fra_{i-1}	Gre_{i-1}	Ita_{i-1}	Ger_{i-2}	Spa_{i-2}	Fra_{i-2}	Gre_{i-2}	Ita_{i-2}	Bel_{i-1}	Ola_{i-1}	Por_{i-1}	USA_{i-1}
Ger_i	0.596	0.475	0.361	0.769	0.301	0.491	0.404	0.313	0.748	0.182	0.338	0.224	0.356	0.065	0.380
Spa_i	0.616	0.451	0.421	0.412	0.476	0.472	0.453	0.403	0.465	0.271	0.651	0.298	0.530	0.096	0.495
Fra_i	0.648	0.528	0.581	0.629	0.361	0.579	0.454	0.496	0.492	0.250	0.682	0.738	0.553	0.218	0.498
Gre_i	0.594	0.409	0.421	0.389	0.329	0.568	0.439	0.311	0.407	0.255	0.332	0.312	0.430	0.088	0.525
Ita_i	0.622	0.430	0.473	0.401	0.552	0.452	0.459	0.511	0.522	0.318	0.590	0.371	0.432	0.082	0.475

con la colonna π rappresentante la probabilità per un elemento della riga di essere generato dalla *slab*. Il modello stimato VAR(2) con priori *spike-and-slab* in $\tau = 0.1$ risulta:

	<i>drift</i>	Ger_{i-1}	Spa_{i-1}	Fra_{i-1}	Gre_{i-1}	Ita_{i-1}	Ger_{i-2}	Spa_{i-2}	Fra_{i-2}	Gre_{i-2}	Ita_{i-2}	Bel_{i-1}	Ola_{i-1}	Por_{i-1}	USA_{i-1}
$Germ_i$	-9.964	-0.063	-0.005	0.006	0.011	0.025	-0.012	-0.014	0.055	0.004	0.057	0.007	-0.024	0.002	0.005
Spa_i	-15.145	0.007	-0.007	-0.010	-0.015	-0.047	-0.004	-0.006	0.006	-0.012	-0.016	0.005	0.002	-0.000	0.029
Fra_i	-13.513	0.009	0.002	0.011	0.002	0.017	0.008	0.017	0.001	-0.001	-0.013	-0.014	0.014	0.002	0.022
Gre_i	-15.782	0.012	0.045	-0.075	-0.001	0.032	0.006	0.002	-0.005	0.007	0.022	0.024	0.039	0.001	0.107
Ita_i	-14.452	-0.019	-0.001	-0.004	-0.052	-0.023	-0.011	-0.000	-0.014	-0.005	-0.016	0.005	0.020	0.001	0.020

mentre le stime per la probabilità a posteriori di appartenere alla componente *slab*, risultano.

	π	Ger_{i-1}	Spa_{i-1}	Fra_{i-1}	Gre_{i-1}	Ita_{i-1}	Ger_{i-2}	Spa_{i-2}	Fra_{i-2}	Gre_{i-2}	Ita_{i-2}	Bel_{i-1}	Ola_{i-1}	Por_{i-1}	USA_{i-1}
$Germ_i$	0.647	0.699	0.434	0.474	0.463	0.576	0.505	0.491	0.723	0.356	0.711	0.412	0.568	0.183	0.516
Spa_i	-0.635	0.483	0.441	0.464	0.533	0.651	0.492	0.407	0.492	0.464	0.530	0.381	0.503	0.159	0.561
Fra_i	0.633	0.472	0.440	0.491	0.394	0.561	0.497	0.516	0.454	0.307	0.503	0.540	0.556	0.187	0.601
Gre_i	-0.668	0.541	0.666	0.752	0.459	0.601	0.478	0.476	0.524	0.446	0.571	0.593	0.647	0.185	0.775
Ita_i	0.636	0.535	0.438	0.399	0.701	0.534	0.500	0.444	0.511	0.356	0.518	0.398	0.536	0.162	0.582

.Le matrici di correlazione stimate nei tre quantili risultano essere:

$\tau = 0.9$	<i>Ger</i>	<i>Spa</i>	<i>Fra</i>	<i>Gre</i>	<i>Ita</i>	$\tau = 0.5$	<i>Ger</i>	<i>Spa</i>	<i>Fra</i>	<i>Gre</i>	<i>Ita</i>	$\tau = 0.1$	<i>Ger</i>	<i>Spa</i>	<i>Fra</i>	<i>Gre</i>	<i>Ita</i>
<i>Ger</i>	17.80	23.02	17.21	11.55	21.84	<i>Ger</i>	10.17	5.49	6.22	3.07	5.68	<i>Ger</i>	16.42	16.03	15.11	13.11	16.50
<i>Spa</i>	23.02	54.72	31.69	21.93	46.18	<i>Spa</i>	5.49	27.66	10.69	4.77	18.96	<i>Spa</i>	16.03	34.47	21.20	17.73	28.04
<i>Fra</i>	17.21	31.69	27.93	14.91	29.36	<i>Fra</i>	6.22	10.69	18.28	3.57	9.85	<i>Fra</i>	15.11	21.20	24.60	15.40	20.95
<i>Gre</i>	11.55	21.93	14.91	21.14	21.56	<i>Gre</i>	3.07	4.77	3.57	17.20	6.08	<i>Gre</i>	13.11	17.73	15.40	23.65	19.00
<i>Ita</i>	21.84	46.18	29.36	21.56	48.02	<i>Ita</i>	5.68	18.96	9.85	6.08	23.92	<i>Ita</i>	6.50	28.04	20.95	19.00	32.96

Tabella 5.7: Stime matrice di covarianza, da Σ .

Sia la stima tramite *spike-and-slab* nei quantili $\tau = (0.1, 0.5)$ sia la stima nel quantile 0.9 sembrano concordare nell'attribuire alla variabile relativa ai prezzi dei *CDS* portoghesi scarso peso nel modello. Sempre guardando le probabilità a posteriori di appartenere alla componente *slab* nei modelli con $\tau = (0.1, 0.5)$, sembra che, tra i termini autoregressivi, i ritardi relativi ai rendimenti spagnoli e greci siano quelli meno importanti. Analizzando le matrici di covarianza stimate si può affermare che le serie siano correlate, sia in mediana che sulle code, con la correlazione tra variabili che è tuttavia molto più forte sulle code. La variabilità stimata per le innovazioni risulta più elevata nel quantile 0.9 rispetto al quantile 0.1. Si può notare una forte correlazione tra Italia e Spagna in tutti e 3 i quantili; le serie delle innovazioni stimate di Spagna e Francia risultano molto correlate nel quantile 0.9. La maggiore correlazione sulle code, osservabile sia dai grafici, figura 5.7, con picchi nei log-rendimenti condivisi da più serie, che dalle matrici di covarianza, è ragionevole alla luce dalle strette relazioni politiche ed economiche che intercorrono tra i paesi analizzati, che quindi possono condividere *shock* comuni tali da generare variazioni più estreme. Viene analizzato, tramite i modelli stimati, il periodo successivo, Gennaio-Luglio 2020, coincidente alla prima ondata di virus Covid-19, per valutare se i modelli per il calcolo del *conditional value at risk* stimati rimangono validi ed utilizzabili anche in questo periodo. L'analisi viene condotta sia graficamente, sia valutando l'adeguatezza del *conditional value at risk* tramite test LRcc. Si può notare già a livello grafico come le serie, figura 5.8, nel periodo Marzo-Aprile, abbiano alcune variazioni estreme, uscendo molteplici volte fuori dagli intervalli di confidenza quantilici stimati. In generale le serie dei log-rendimenti sembrano avere, nel periodo in questione, un comportamento molto simile tra loro. I test LRcc eseguiti rifiutano l'ipotesi di corretta stima del quantile 0.9 da

parte del modello stimato, indicando quindi che il comportamento della serie in questo periodo, nel quantile 0.9, è diverso da quello del periodo precedente. I modelli stimati sembrano invece stimare correttamente il comportamento della serie nei quantili 0.1 e 0.5.

quantile	Germania	Spagna	Francia	Grecia	Italia	LRcc	p-value
0.1	0.107	0.143	0.107	0.071	0.071	4.197	0.123
0.5	0.536	0.571	0.679	0.429	0.500	2.112	0.348
0.9	0.857	0.821	0.857	0.714	0.750	25.158	0.000

Viene riportata anche qui la stima statica del *value at risk* costante effettuata utilizzando i quantili stimati nel periodo di stima.

quantile	Germania	Spagna	Francia	Grecia	Italia	LRcc	p-value
0.1	0.207	0.276	0.241	0.241	0.172	21.296	0.000
0.5	0.552	0.586	0.655	0.517	0.517	5.020	0.081
0.9	0.862	0.827	0.862	0.689	0.758	24.876	0.000

Sia il *value at risk* statico che il *conditional value at risk* ottenuto tramite modello VAR quantilico performano male nel quantile 0.9, probabilmente per la forte crescita dei prezzi dei CDS avvenuta nel periodo di Marzo-Aprile. I modelli stimati nei quantili 0.5, 0.1, a differenza del *value at risk* statico, risultano invece statisticamente validi ed applicabili per il calcolo del *conditional value at risk* a ritardo 1.

5.4.2 Analisi strutturali

Per comprendere al meglio le dinamiche del sistema vengono svolte alcune analisi strutturali quantiliche. La prima analisi, volta a valutare l'apporto alla varianza delle innovazioni del sistema da parte delle variabili modellate, viene effettuata stimando le *generalised quantile forecast error variace decomposition* per i 3 quantili. Anche in questo caso, trattandosi sempre di dati settimanali, si è scelto un orizzonte temporale di 4 settimane. La stima della GQFEVD per $\tau = 0.5$ risulta:

	<i>Ger</i>	<i>Spa</i>	<i>Fra</i>	<i>Gre</i>	<i>Ita</i>
<i>Ger</i>	0.379	0.322	0.549	0.301	0.356
<i>Spa</i>	0.200	0.213	0.143	0.271	0.204
<i>Fra</i>	0.187	0.207	0.149	0.118	0.167
<i>Gre</i>	0.136	0.147	0.101	0.125	0.191
<i>Ita</i>	0.098	0.111	0.058	0.186	0.083

la stima delle GQFEVD per $\tau = 0.1$ risulta:

	<i>Ger</i>	<i>Spa</i>	<i>Fra</i>	<i>Gre</i>	<i>Ita</i>
<i>Ger</i>	0.572	0.491	0.481	0.491	0.512
<i>Spa</i>	0.135	0.127	0.114	0.184	0.206
<i>Fra</i>	0.093	0.090	0.092	0.069	0.095
<i>Gre</i>	0.118	0.222	0.270	0.080	0.124
<i>Ita</i>	0.083	0.070	0.043	0.175	0.062

la stima delle GQFEVD per $\tau = 0.9$ risulta:

	<i>Ger</i>	<i>Spa</i>	<i>Fra</i>	<i>Gre</i>	<i>Ita</i>
<i>Ger</i>	0.684	0.674	0.654	0.653	0.692
<i>Spa</i>	0.158	0.154	0.150	0.177	0.140
<i>Fra</i>	0.057	0.054	0.065	0.048	0.063
<i>Gre</i>	0.054	0.074	0.092	0.061	0.073
<i>Ita</i>	0.048	0.044	0.039	0.061	0.032

Si può subito notare come, la variabile che più di tutte sembra influenzare la varianza delle innovazioni del processo, sia il log-rendimento dei *CDS* tedeschi. In particolare, al crescere di τ , quindi all'avvicinarsi ad una condizione negativa per i titoli di stato, con l'aumento del rischio associato, sembra che la variabile tedesca diventi man mano più importante nel determinare le variazioni del sistema. Viene successivamente svolta una analisi tramite *generalised quantile impulse response function* per valutare se, rispetto a quanto visto nell'analisi pre-crisi, il sistema dei *CDS* risulti più stabile ad impulsi sui log-rendimenti greci. Viene mostrato a livello grafico l'effetto di uno shock

sulla variabile greca, e l'impatto che ha nell'arco delle 8 settimane successive. Vengono mostrati i risultati in figura 5.10. Appare evidente come, nei quantili $\tau = 0.1$ e $\tau = 0.5$, l'impatto che un impulso sulla variabile greca ha sul sistema risulti sostanzialmente nullo. Il quantile $\tau = 0.9$ risulta essere l'unico in cui, un eventuale shock dei rendimenti dei *CDS* greci, sembri influenzare, seppur leggermente, il sistema. Il quantile 0.9 dei log-rendimenti dei *CDS* italiani sembra essere l'unico realmente influenzato da un eventuale shock dei *CDS* greci. Sembra quindi ragionevole affermare che, secondo i modelli utilizzati, il sistema dei *CDS* analizzati risulti più solido e resistente a shock sui *CDS* greci, rispetto al periodo precedentemente analizzato. Data la rilevanza che i log-rendimenti tedeschi hanno mostrato nelle *generalised quantile forecast error variace decomposition* viene effettuato una analisi tramite *generalised quantile impulse response function* con impulso sulla variabile tedesca, figura 5.9, mostrando le conseguenze sul sistema nelle 8 settimane successive. Anche in questo caso si ha che l'unica variazione significativa si ha sul quantile 0.9 dei log-rendimenti dei *CDS* italiani. Per completare l'analisi del sistema è stato effettuato uno stress test sul sistema. Rispetto a quanto fatto nella precedente sezione, in cui lo shock introdotto è stato di tipo asimmetrico sulla variabile greca, in questo caso è stata simulata una situazione di stress direttamente su tutto il sistema. Sono state generate 6 osservazioni di fila sul quantile 0.9, generate tramite il modello stimato nel quantile, seguite da 6 osservazioni generate in mediana con il modello stimato nel quantile 0.5. I risultati dello *stress test* vengono riportati in figura 5.11. Rispetto allo *stress test* effettuato sul sistema pre-crisi il sistema risulta molto più stabile, con una crescita del prezzo meno esplosiva nelle settimane di stress, seguite poi da 6 settimane in cui il prezzo si stabilizza e inizia una decrescita leggera.

5.5 Conclusione dell'analisi

I metodi utilizzati si sono rivelati utili ed applicabili all'analisi delle serie dei prezzi dei *CDS*, trasformati in log-rendimenti, analizzata. I modelli VAR quantilici proposti hanno dimostrato un sufficiente adattamento ai dati nell'insieme di stima, con performance simili a quelle di un modello stimato equazione per equazione. Nel caso dell'analisi pre-crisi il modello proposto

per la stima della distribuzione nei 3 quantili si è dimostrato adeguato per la stima del *conditional value at risk*, con performance tuttavia simili a quelle che si otterrebbero tramite un semplice *value at risk* basato sui quantili dell'insieme di stima. L'utilizzo di metodi di analisi strutturale quantilica è stato in grado di fornire utili informazioni sul sistema dei *CDS* analizzati pre-crisi, evidenziando l'importanza dei *CDS* greci nel periodo nel determinare le innovazioni del modello, e l'impatto di un impulso sulla variabile greca. Lo *stress test*, valutato simulando una serie di shock successivi sulla variabile greca, ha generato una situazione molto simile a quella poi osservata nel corso dell'anno successivo, mostrando quindi che, tramite un'opportuna modellazione e l'uso di opportuni *stress test*, si potessero già rilevare le criticità di sistema poi osservate. L'analisi del periodo post-crisi, svolta tramite modelli leggermente più complicati, ha mostrato come, rispetto al periodo pre-crisi, il sistema fosse più stabile a shock sulle singole variabili, specie nei quantili 0.1, 0.5. Il quantile 0.9, relativo a shock positivi sui *CDS*, equivalente ad un aumento di rischio dei titoli di stato del paese di riferimento, è l'unico a mostrare una certa reattività agli impulsi. I rendimenti tedeschi sembrano avere un ruolo di predominanza nel determinare le innovazioni del sistema in questo periodo. Le *impulse response function* con impulsi su Grecia e Germania non hanno rivelato particolare reattività nel sistema, se non per la variabile Italia nel quantile 0.9, quantile associato ad un aumento del rischio percepito sui titoli di stato. Il calcolo di uno *stress test* simmetrico restituisce una serie molto più stabile rispetto a quella pre-crisi, con una crescita dei prezzi minore nonostante l'intero sistema venga posto in una situazione di stress.

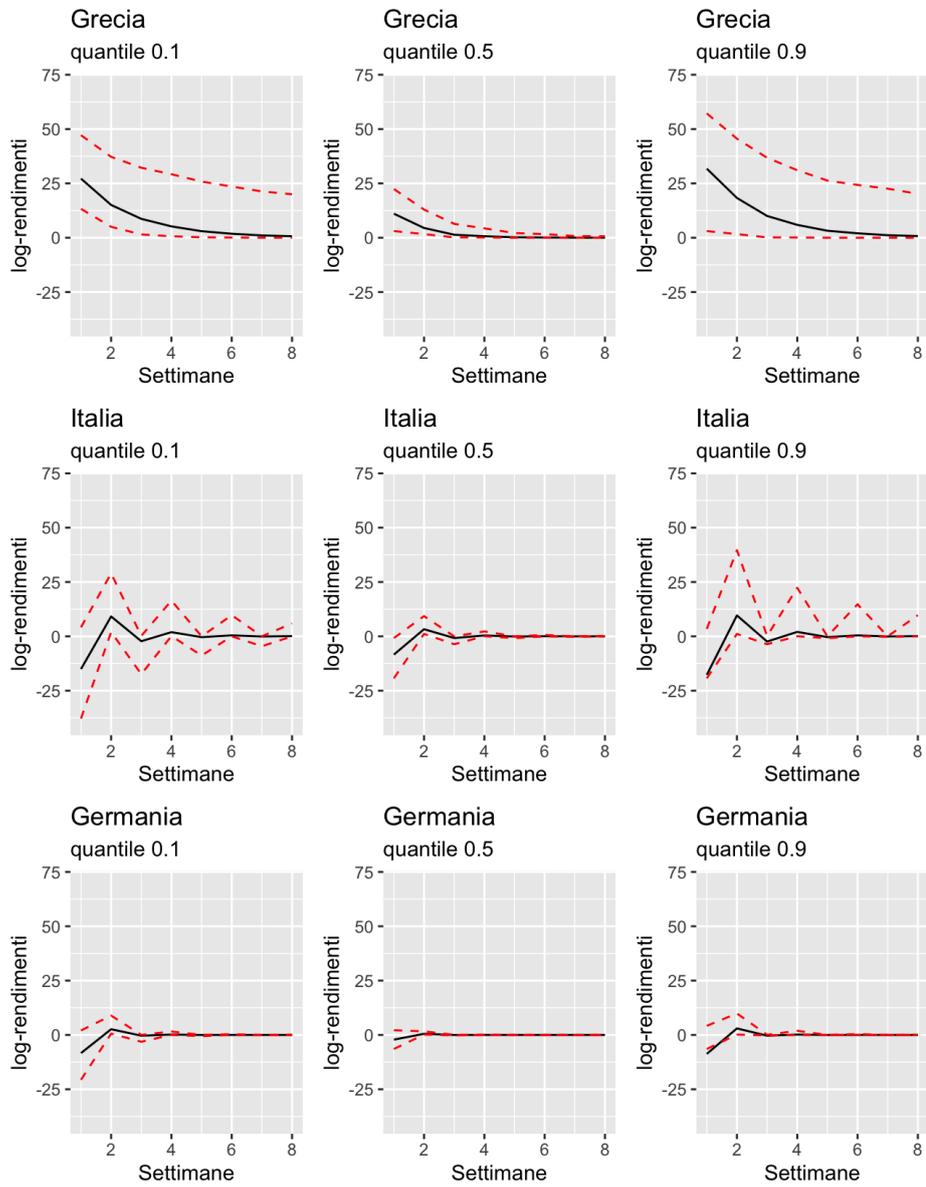


Figura 5.5: Impulse response function a 8 settimane per uno shock sulla variabile Grecia.

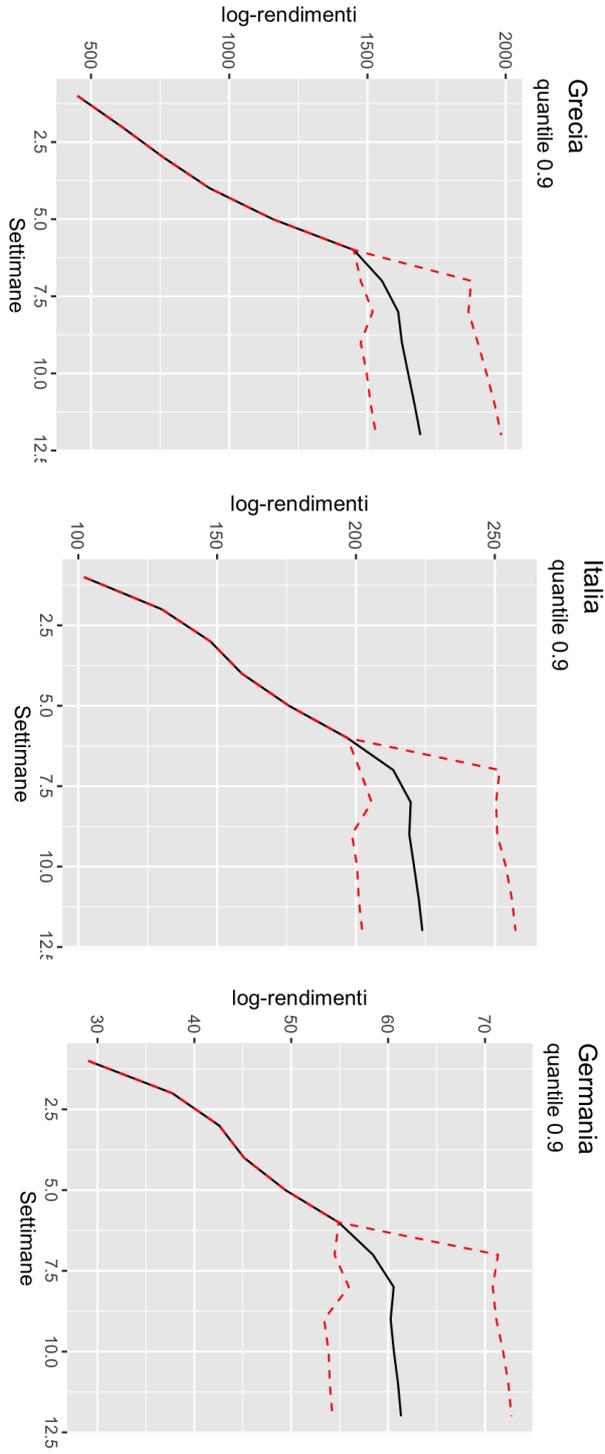


Figura 5.6: Stress test a 12 settimane sui prezzi dei CDS di Grecia, Italia e Germania.

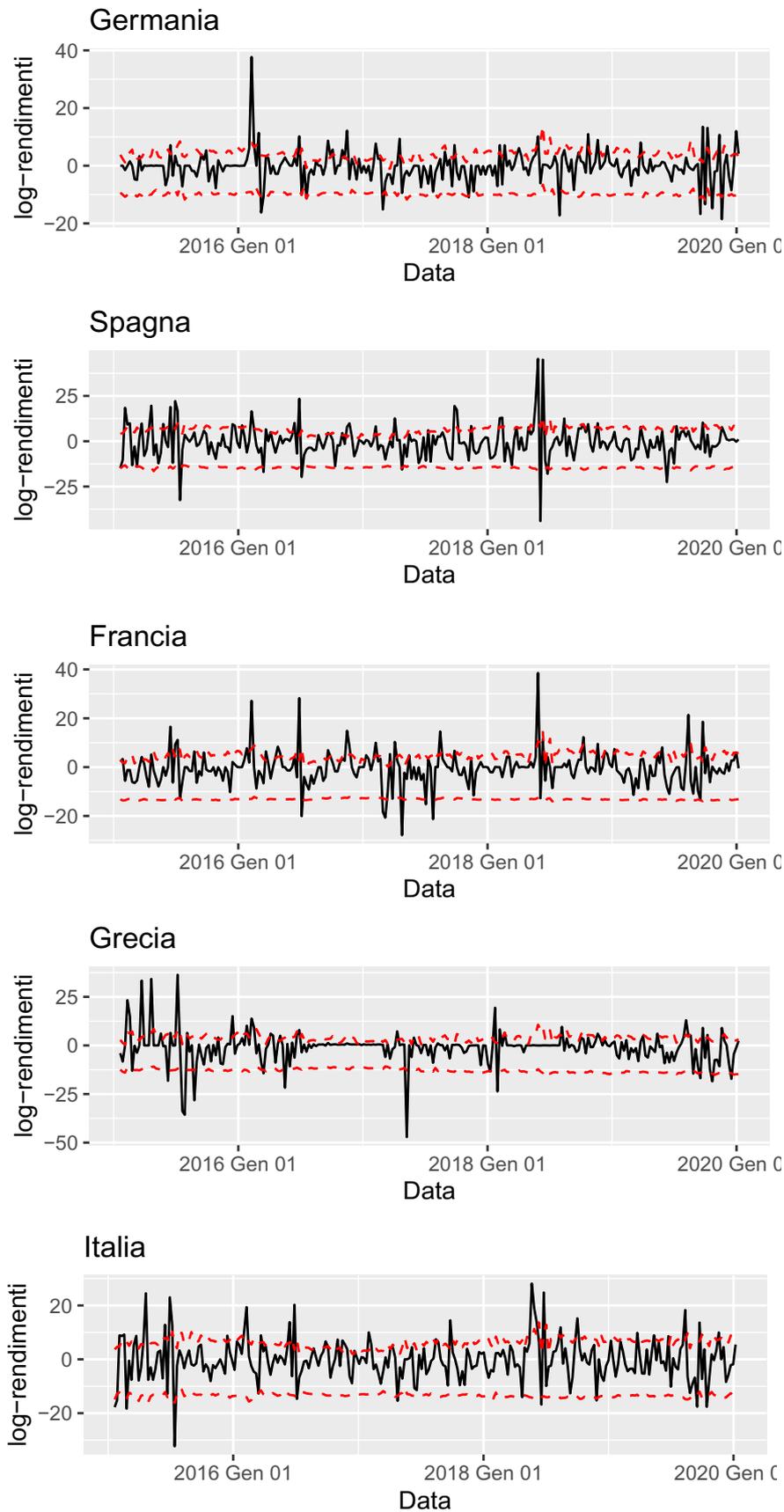


Figura 5.7: Serie dei prezzi dei CDS, 2015-2019, con relativi quantili 0.9, 0.1 stimati.

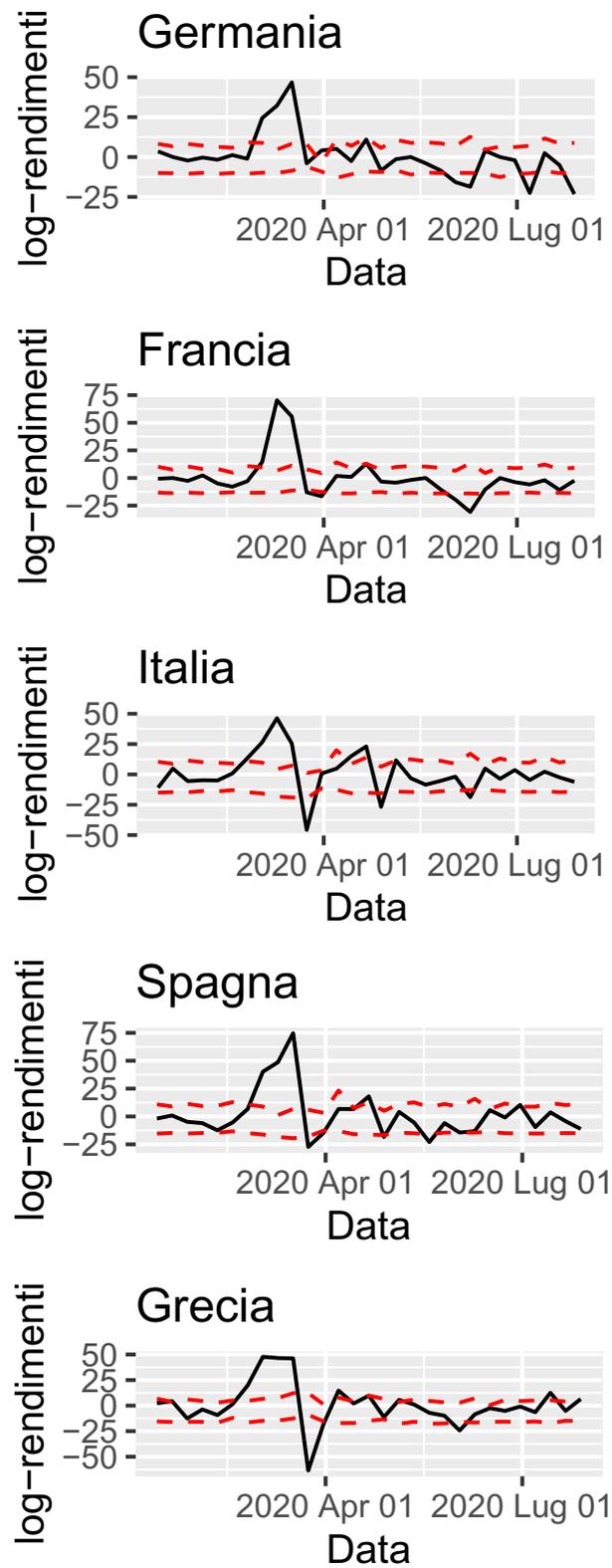


Figura 5.8: Confronto quantili previsti, tramite modelli VAR stimati nei quantili 0.1, 0.9, e serie osservata dei log-rendimenti dei *CDS* per il 2020.

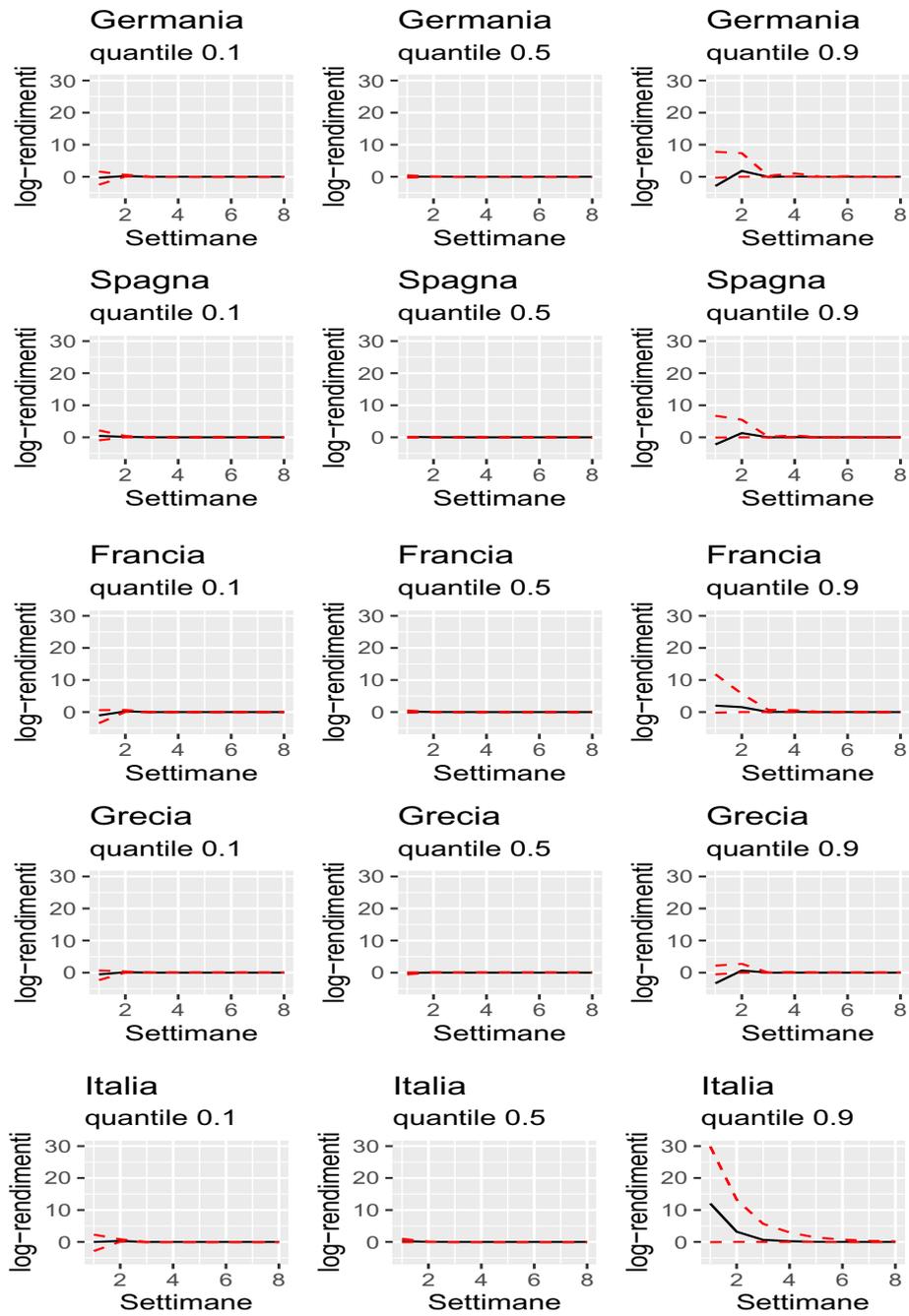


Figura 5.9: Generalised quantile IRF a 8 settimane per uno shock sui *CDS* tedeschi.

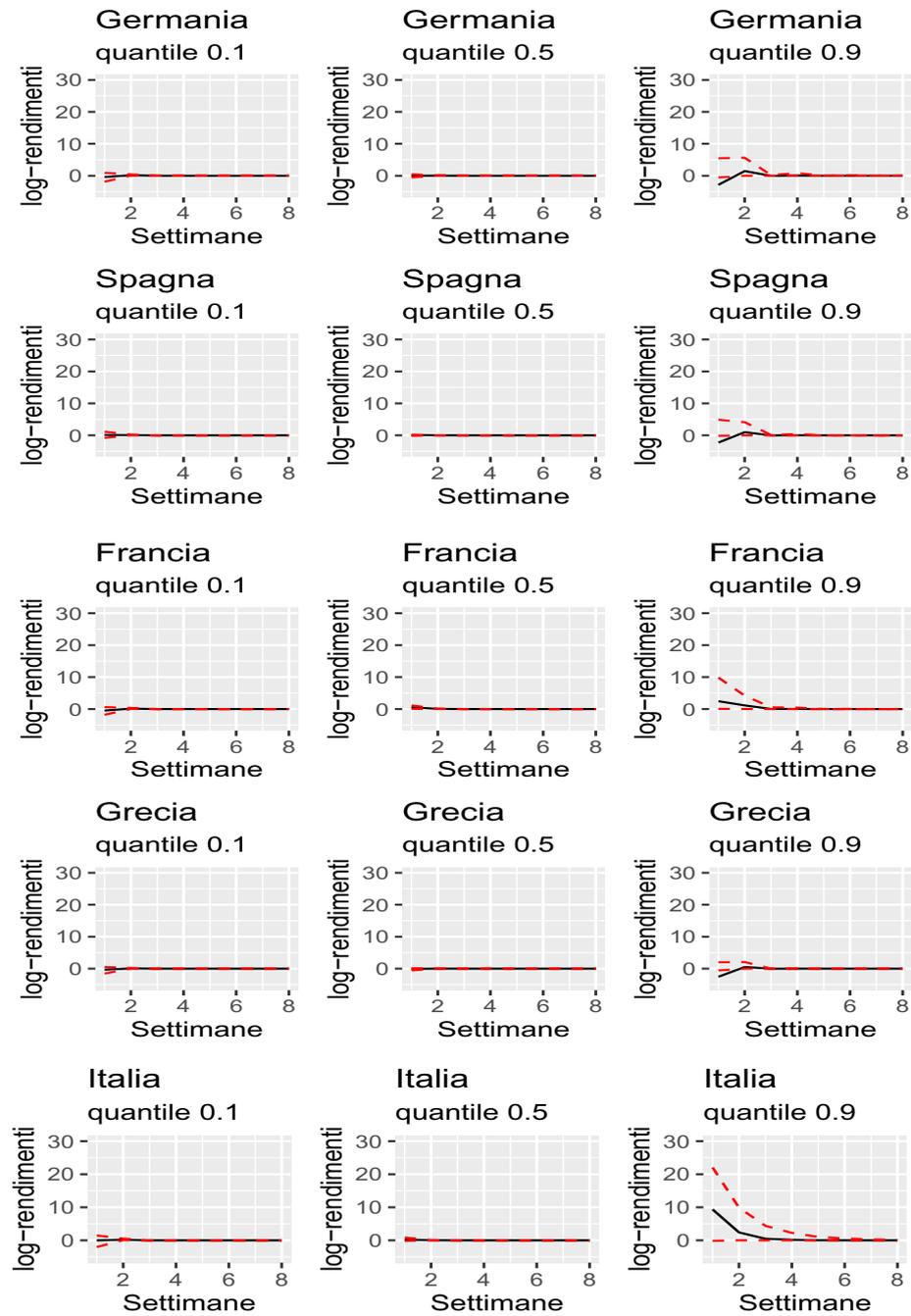


Figura 5.10: Generalised quantile IRF a 8 settimane per uno shock sui *CDS* greci.

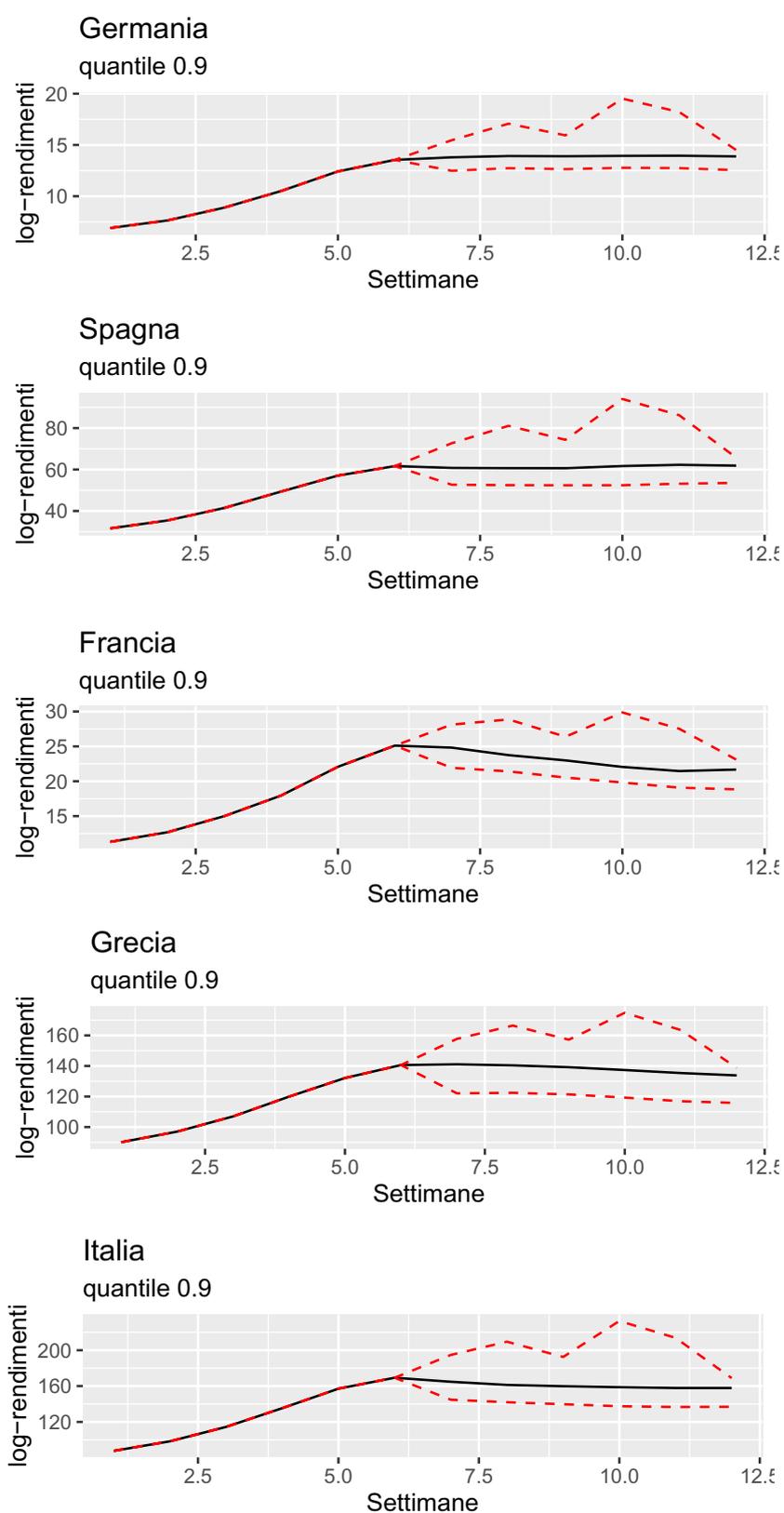


Figura 5.11: Stress test simmetrico sui prezzi dei *CDS* di Germania, Spagna, Francia, Grecia e Italia.

Conclusioni

La modellazione VAR quantilica è un valido complemento alla modellazione VAR per la media condizionata, fornendo una maggiore conoscenza della distribuzione del fenomeno di interesse. In questa tesi ho cercato di ampliare, sfruttando la maggiore flessibilità data dalla modellazione bayesiana, le possibilità a disposizione, fornendo alcune alternative per la stima di un VAR quantilico e proponendo alcune estensioni. Le simulazioni nel capitolo 4 hanno mostrato le ottime capacità dei modelli proposti di stimare correttamente il modello VAR generatore. L'estensione *spike-and-slab* per il modello VAR quantilico si è dimostrata efficace nelle simulazioni effettuate nel capitolo 4, ed ha ottenuto degli ottimi risultati nella modellazione del *conditional value at risk* nel capitolo 5, fornendo anche informazioni sull'importanza delle variabili nel modello. Il modello VAR quantilico con restrizioni ha mostrato nelle simulazioni di poter fornire ottimi risultati sotto corretta specificazione delle restrizioni; in futuro potrebbe risultare di interesse valutare la stima di un modello con restrizioni e vincoli anche sulla matrice di covarianza. Gli strumenti di analisi strutturale quantilica si sono dimostrati un utile strumento per comprendere meglio i modelli stimati; rispetto alle tecniche di analisi strutturale finora proposte in letteratura per i VAR quantilici, basate sulla impostazione frequentista, le tecniche proposte presentano il vantaggio di poter fornire intervalli di probabilità per le quantità generate, utilizzabili anche per test statistici, oltre ad essere direttamente derivabili dalle quantità stimate tramite MCMC. Gli algoritmi proposti hanno dimostrato ottime performance sia sui dati che nelle simulazioni, non risultando tuttavia adeguati per la stima di VAR ad elevata dimensionalità, come visto per le simulazioni con variabile risposta in dimensione $d = 10$. Si lascia ad eventuali sviluppi

futuri l'implementazione di una versione degli algoritmi e modelli proposti adeguata al caso di modelli autoregressivi vettoriali quantilici in dimensione maggiore. L'applicazione della modellazione quantilica per la stima del *conditional value at risk* nel capitolo 5 ha ottenuto risultati efficaci, adattandosi bene alle serie analizzate, oltre a mostrare una discreta capacità di previsione anche per periodi successivi a quelli in cui è stata stimata la serie, e fornendo interessanti spunti sulle caratteristiche del sistema e sui rischi presenti, tramite l'applicazione di *quantile impulse response function*, *quantile forecast error variance decomposition* e l'utilizzo di procedure di *stress test* che sfruttano i modelli VAR quantilici stimati. La modellazione bayesiana per i modelli VAR quantilici offre una maggiore flessibilità per la creazione di modelli, flessibilità che potrà essere sfruttata in futuro per la creazione di estensioni e nuovi strumenti per la previsione di quantili estremi.

Bibliografia

- Adrian, T. e M. Brunnermeier. «CoVaR». In: *American Economic Review* 106.7 (lug. 2016), pp. 1705–41. DOI: [10.1257/aer.20120555](https://doi.org/10.1257/aer.20120555). URL: <https://www.aeaweb.org/articles?id=10.1257/aer.20120555>.
- Andrieu, C. et al. «An Introduction to MCMC for Machine Learning». In: *Machine Learning* 50 (gen. 2003), pp. 5–43. DOI: [10.1023/A:1020281327116](https://doi.org/10.1023/A:1020281327116).
- Bassett, G. e R. Koenker. «Regression quantiles». In: *Econometrica* 46.1 (1978), pp. 33–50.
- Cecchetti, S. e L. Li. «Do capital adequacy requirements matter for monetary policy?» In: *Economic Inquiry* 46.4 (2008), pp. 643–659.
- Chavleishvili, S. e S. Manganelli. «Forecasting and stress testing with quantile vector autoregression». In: (2019).
- Christoffersen, P. F. «Evaluating Interval Forecasts». In: *International Economic Review* 39.4 (1998), pp. 841–862. ISSN: 00206598, 14682354. URL: <http://www.jstor.org/stable/2527341>.
- David, A, B. Kris e L. Catania. «Generalized Autoregressive Score Models in R: The GAS Package». In: *Journal of Statistical Software* 88.6 (2019), pp. 1–28. DOI: [10.18637/jss.v088.i06](https://doi.org/10.18637/jss.v088.i06).
- Eikon., Thomson Reuters. «Sovereign Credit Default Swap». In: *Journal of Statistical Software* (2018). URL: <https://eikon.thomsonreuters.com/index.html>.
- Fatti, L. P., E. M. Senoana e M. L. Thompson. «Bayesian updating in reference centile charts». In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 161.1 (1998), pp. 103–115.
- Hao, L. e D. Naiman. *Quantile regression*. 149. Sage, 2007.

- Kaufmann, S. e C. Schumacher. *Bayesian estimation of sparse dynamic factor models with order-independent identification*. Working Papers 13.04. Swiss National Bank, Study Center Gerzensee, 2013. URL: <https://EconPapers.repec.org/RePEc:szg:worpap:1304>.
- Korobilis, D. «Quantile forecasts of inflation under model uncertainty». In: *Available at SSRN 2610253* (2015).
- Kotz, S., T. J Kozubowski e K. Podgórski. «Asymmetric multivariate Laplace distribution». In: *The Laplace distribution and generalizations*. Springer, 2001, pp. 239–272.
- Kozumi, H. e G. Kobayashi. «Gibbs sampling methods for Bayesian quantile regression». In: *Journal of statistical computation and simulation* 81.11 (2011), pp. 1565–1578.
- Kruschke, L. «Bayesian data analysis for newcomers». In: *Psychon Bull Review* 25 (2018), pp. 155–177.
- Kupiec, P. *Techniques for verifying the accuracy of risk measurement models*. Finance and Economics Discussion Series 95-24. Board of Governors of the Federal Reserve System (U.S.), 1995. URL: <https://EconPapers.repec.org/RePEc:fip:fedgfe:95-24>.
- «Techniques for verifying the accuracy of risk measurement models». In: *The J. of Derivatives* 3.2 (1995).
- Li, G., Y. Li e C. Tsai. «Quantile correlations and quantile autoregressive modeling». In: *Journal of the American Statistical Association* 110.509 (2015), pp. 246–261.
- Lütkepohl, H. *New introduction to multiple time series analysis*. Berlin [u.a.]: Springer, 2005. XXI, 764.
- Naifar, N. «What Explains the Sovereign Credit Default Swap Spreads Changes in the GCC Region?» In: *Journal of Risk and Financial Management* 13.10 (2020), p. 245.
- Pesaran, H. e Y. Shin. «Generalized impulse response analysis in linear multivariate models». In: *Economics letters* 58.1 (1998), pp. 17–29.
- Ročková, V. e E. George. «EMVS: The EM approach to Bayesian variable selection». In: *Journal of the American Statistical Association* 109.506 (2014), pp. 828–846.

-
- Schüler, Y. «Asymmetric effects of uncertainty over the business cycle: A quantile structural vector autoregressive approach». In: (2014).
- White, H., T. Kim e S. Manganelli. «VAR for VaR: Measuring tail dependence using multivariate regression quantiles». In: *Journal of Econometrics* 187.1 (2015), pp. 169–188.
- Yu, K. e R. Moyeed. «Bayesian quantile regression». In: *Statistics & Probability Letters* 54.4 (2001), pp. 437–447.

Appendice A

Capitolo 5

Analisi Grecia, Italia, Germania

Point est.	Upper C.I.	Point est.	Upper C.I.	Point est.	Upper C.I.
1.26	2.00	1.03	1.04	1.01	1.02
1.31	2.37	1.18	1.19	1.20	1.23
1.25	1.70	1.05	1.05	1.22	1.38
1.24	1.51	1.15	1.16	1.01	1.01
1.24	1.57	1.03	1.03	1.15	1.17
1.11	1.20	1.00	1.00	1.14	1.20
1.29	1.63	1.08	1.09	1.00	1.01
1.12	1.14	1.13	1.15	1.02	1.03
1.22	1.33	1.00	1.00	1.02	1.02
1.12	1.19	1.11	1.11	1.00	1.01
1.28	1.60	1.22	1.24	1.05	1.06
1.28	1.55	1.03	1.05	1.03	1.04
1.09	1.23	1.06	1.06	1.02	1.02
1.10	1.17	1.20	1.20	1.04	1.05
1.13	1.20	1.00	1.00	1.12	1.13
1.21	1.42	1.14	1.15	1.00	1.00
1.24	1.26	1.22	1.24	1.04	1.05
1.18	1.21	1.02	1.02	1.02	1.02
1.28	1.70	1.16	1.18	1.04	1.06
1.22	1.42	1.22	1.24	1.02	1.03
1.13	1.19	1.01	1.02	1.15	1.22

Tabella 8: *Potential scale reduction factor per β nei quantili 0.1, 0.5, 0.9.*

693.26	3061.15	480.50
736.03	3239.13	514.81
703.25	4173.96	389.53
1297.90	2577.29	771.47
1612.89	3427.67	977.99
1544.94	2487.38	970.20
797.67	3000.00	1374.79
763.77	3365.21	1269.42
712.11	3300.18	1436.63
1071.79	3670.58	960.51
942.67	4486.10	903.10
844.67	3054.33	1327.66
1648.22	3210.99	735.72
845.19	3160.62	782.83
1149.76	3023.38	669.34
719.93	2962.42	1600.16
874.81	4231.83	1436.98
1085.32	3279.87	1252.50
939.82	2970.55	1541.37
931.84	6152.78	1406.86
1081.71	2878.72	1469.13

Tabella 9: Effective sample size per β , quantili 0.1,0.5,0.9.

Analisi post-crisi

quantile	β	Ω	D	v	π	σ_ε	K_1	μ
0.1	1.05	1.01	1.07	1.12	1.01	1.01	1.1	1.02
0.5	1.10	1.02	1.04	1.17	1.01	1.01	1	1.03
0.9	1.09	1.03	1.05	1.12	-	-	-	-

Tabella 10: *Potential scale reduction factors* multivariato per le catene generate nei modelli selezionati.

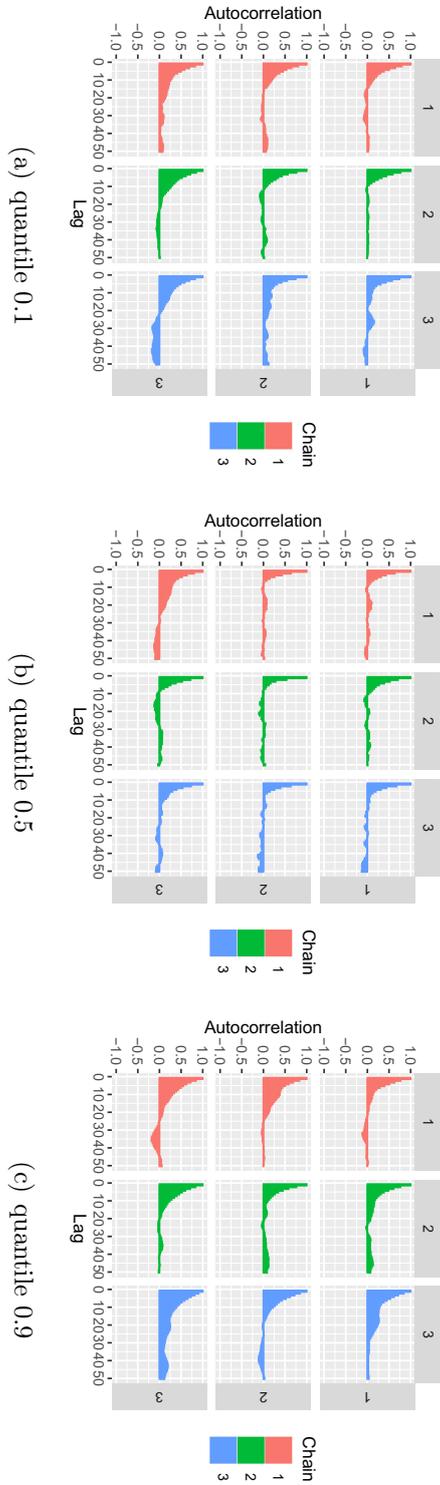
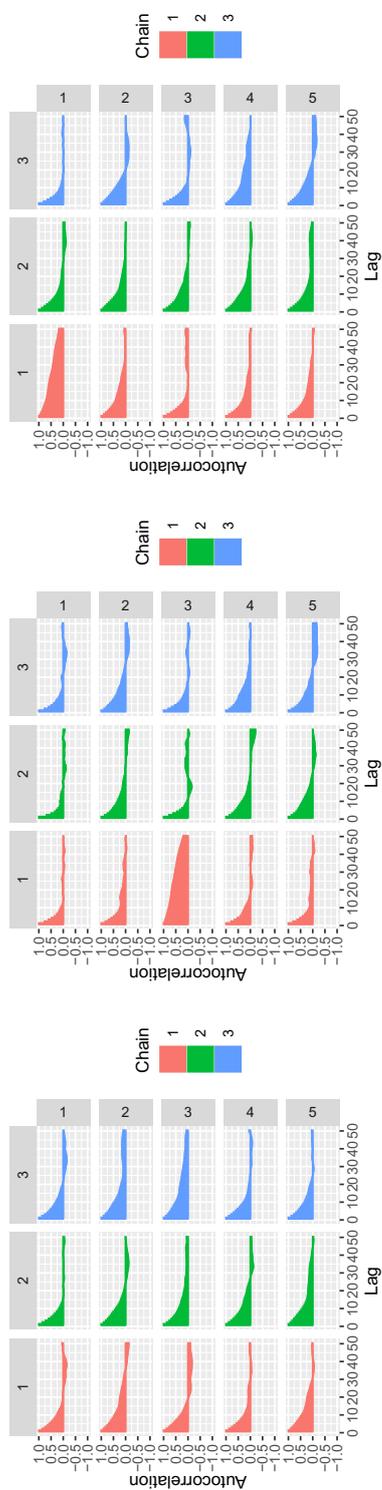


Figura 12: Autocorrelazione per le catene di D generate nei 3 quantili, per il modello stimato pre-crisi.



(a) quantile 0.1

(b) quantile 0.5

(c) quantile 0.9

Figura 13: Autocorrelazione per le catene di D generate nei 3 quantili, per il modello stimato post-crisi.

quantile	β	Ω	D	v	π	σ_ε	K_1	μ
0.1	2333	828	307	1066	2817	2236	2201	618
0.5	2536	2044	287	1736	3448	2709	2739	525
0.9	1853	1347	306	1102	-	-	-	-

Tabella 11: *Effective sample size* media per le catene generate nei modelli selezionati.

Appendice B

Vengono qui riportati i codici per gli algoritmi mostrati nel capitolo 4. Tutti i codici riportati sono stati scritti in R. Il primo codice riportato è relativo alla stima dell'algoritmo presentato in 4.2.1.

Codice 1: *Algoritmo stima VAR quantilico*

```
require(ks)
require(GIGrvg)
require(LaplacesDemon)
require(SimDesign)
require(Matrix)
require(corpcor)
require(utils)

####posteriori per D
lpost_inomega=function(omega, resvec, resomega, cn, COVi){
  invomega=solve(omega)
  primotermine=vec(invomega%*%t(resomega))*resvec
  log(det(COVi))*-(cn)/2-1/2*sum(primotermine)
}

###funzione per generare la matrice delle covariate
createZ=function(y,x,p,d,t,m){
  #p is the retard we want in the model
  Z=matrix(NA,d*p+1+m,t-p)
  Z[1,]=rep(1,t-p)
  if(p>0){
    for(j in 1:p){
```

```

        Z[(2+d*(j-1)):(d*j+1),]=t(y[(p-j+1):(t-j),])
    }}
  if(m>1){
    Z[(2+d*p):(d*p+m-1+2),]=t(x[(1+p):t,])
  }
  if(m==1){
    Z[(2+d*p):(d*p+m-1+2),]=(x[(1+p):t])
  }
  Z
}
## algoritmo di stima
mcmc1=function(Y,X=NULL,p,q,R,burnin=5000,c){
  #dimensioni
  d=dim(Y)[2]#t*d
  t=dim(Y)[1]
  m=dim(X)[2]#t*m
  if(is.null(X)){
    m=0
  }
  if(is.null(m)){
    m=1
  }

  #preparo le mie y e x
  Z=createZ(Y,X,p,d,t,m)

  #tolgo le prime p osservazioni
  Y=Y[(p+1):t,]
  y=vec(t(Y))

  #costanti funzione del quantile q
  lambda=(1-2*q)/(q*(1-q))
  sigma2=sqrt(2/(q*(1-q)))

  #quantita ripetute nel gibbs
  pargig1=(2-d)/2
  #C=diag(rep(10,d))+matrix(5,d,d)
  #elemnti utili gibbs
  acc=0

```

```

##adesso creo le matrici in cui mettere i miei risultati
#da qui in poi uso t1 al posto di t
t1=t-p
v=matrix(NA,t1,R)
beta=matrix(NA,d*d*p+d*d*m,R)
SIGMA=array(NA,dim=c(d,d,R))
sqD=matrix(NA,d,R)

#prima iterazione
v[,1]=rexp(t1)
beta[,1]=runif(dd1,-1,1)
Ri=cov2cor(cov(Y))
SIGMA[,1]=diag(sigma2)%*%Ri%*%diag(sigma2)
sqD[,1]=runif(d,0,5)
VARi=diag(sqD[,1])%*%SIGMA[,1]%*%diag(sqD[,1])
cn=t-d
#mcmc
pb=txtProgressBar(0,R,initial=0)
soglie=runif(R)
for(i in 2:R){
  if(i%50==0){setTxtProgressBar(pb,i)}
  ##genero per primo beta
  eps_tilde=Y-matrix(v[,i-1],t1,1)%*% (matrix(sqD[,i-1],1,d)*lambda)
  ZV=Z%*%diag(1/v[,i-1])
  ZVZ=ZV%*%t(Z)
  inv_ZVZ=solve(ZVZ/2+t(ZVZ)/2)

  #par.posteriori beta
  sigmabpost=kronecker(inv_ZVZ,VARi)
  sigmabpost=(sigmabpost/2+t(sigmabpost)/2)
  mubpost=vec(t(eps_tilde)%*%diag(1/v[,i-1])%*%t(Z)%*%t(inv_ZVZ))
  beta[,i]=rmvnorm(1,mubpost,sigmabpost)
  Bi=matrix(beta[,i],d*p+1+m,d,byrow=T)
  #####
  #genero i fattori latenti

```

```

invVARi=pseudoinverse(VARi)
pargig3=2+t(sqD[,i-1])%%diag(lambda)%%invVARi%%diag(lambda)%%sqD[,
  i-1]
eps=Y-t(Z)%%Bi

for(j in 1:t1){
  pargig2=t(eps[j,])%%invVARi%%eps[j,]
  v[j,i]=rgig(1,pargig1,pargig2,pargig3)
}
#####
#genero D, tramite mh

sqDstar=sqD[,i-1]+c*runif(d,-1,1)
sqDstar=maxd(sqDstar)
VARistar=diag(sqDstar)%%SIGMA[,i-1]%%diag(sqDstar)

Eps_tot_star=Y-t(Z)%%Bi-matrix(v[,i],t1,1)%% (matrix(sqDstar,1,d)*
  lambda)
Eps_omega_star=Eps_tot_star/sqrt(v[,i])

Eps_tot=Y-t(Z)%%Bi-matrix(v[,i],t1,1)%% (matrix(sqD[,i-1],1,d)*
  lambda)
Eps_omega=Eps_tot/sqrt(v[,i])
#passo metropolis

COVi=diag(sqD[,i-1])%%Ri%%diag(sqD[,i-1])
COVstar=diag(sqDstar)%%Ri%%diag(sqDstar)
alpha=exp(lpost_inomega(VARistar,vec(t(Eps_omega_star)),
Eps_omega_star,cn,COVstar)-lpost_inomega(
VARi,vec(t(Eps_omega)),Eps_omega,cn,COVi))
sqD[,i]=sqD[,i-1]

if(soglie[i]<alpha){
  sqD[,i]=sqDstar
  acc=acc+1
}

```

```
#####
#genero Omega
Eps_tot=Y-t(Z)%*%Bi-matrix(v[,i],t1,1)%*% (matrix(sqD[,i],1,d)*lambda)
SIGMA_post=diag(1/sigma2)%*%diag(1/sqD[,i])%*%
t(Eps_tot)%*%diag(1/v[,i])%*%Eps_tot%*%
diag(1/sqD[,i])%*%diag(1/sigma2)
Ri=cov2cor(diag(1/sigma2)%*%
rinvwishart(t1,SIGMA_post)%*%diag(1/sigma2))
SIGMA[,i]=diag(sigma2)%*%Ri%*%diag(sigma2)
VARi=diag(sqDstar)%*%SIGMA[,i]%*%diag(sqDstar)
}
#ritorno i valori
beta=beta[,burnin:R]
sigma=SIGMA[,burnin:R]
latent_factor=v[,burnin:R]
sqD=sqD[,burnin:R]
MM=seq(1,R-burnin,by=10)
beta=beta[,MM]
sigma=sigma[,MM]
latent_factor=v[,MM]
sqD=sqD[,MM]
return(list(accepted=acc/R,beta=beta,
sigma=sigma,latent_factor=latent_factor,
sqD=sqD,quantile=q,y=y,x=Z,soglia=alpha))}
```

Viene riportato di seguito il codice relativo alla stima dell'algorithmo presentato in 4.2.2.

Codice 2: *Algorithmo stima VAR con generazione Ω_d*

```
#algorithmo di stima vero e proprio
mcmc1=function(Y,X=NULL,p,q,dd1,R,burnin=5000){
  #dimensioni
  d=dim(Y)[2]#t*d
  t=dim(Y)[1]
  m=dim(X)[2]#t*m
```

```
if(is.null(X)){
  m=0
}

#preparo le mie y e x
Z=createZ(Y,X,p,d,t,m)

#tolgo le prime p osservazioni
Y=Y[(p+1):t,]
y=vec(t(Y))

#costanti funzione del quantile q
lambda=(1-2*q)/(q*(1-q))
sigma2=2/(q*(1-q))

#quantita ripetute nel gibbs
pargig1=(2-d)/2
#elemnti utili gibbs
soglie=runif(R)
acc=0
alpha=rep(0,R)
##adesso creo le matrici in cui portero i miei risultati
#da qui in poi uso t1 al posto di t
t1=t-p
v=matrix(NA,t1,R)
beta=matrix(NA,dd1,R)
SIGMA=array(NA,dim=c(d,d,R))

#prima iterazione
v[,1]=rexp(t1)
beta[,1]=runif(dd1,-1,1)
SIGMA[, ,1]=cov(Y)/(d*sigma2)
cn=t
#mcmc
pb=txtProgressBar(0,R,initial=0)
```

```

for(i in 2:R){
  if(i%%50==0){setTxtProgressBar(pb,i)}
  ##genero per primo beta
  eps_tilde=Y-lambda*matrix(v[,i-1],t1,1)%*% sqrt(diag(SIGMA[, , i-1]))
  ZV=Z%*%diag(1/v[,i-1])
  ZVZ=ZV%*%t(Z)
  inv_ZVZ=solve(ZVZ/2+t(ZVZ)/2)

  #par.posteriori
  sigmabpost=sigma2*kronecker(inv_ZVZ,SIGMA[, , i-1])
  sigmabpost=(sigmabpost/2+t(sigmabpost)/2)
  mubpost=vec(t(eps_tilde)%*%diag(1/v[,i-1])%*%t(Z)%*%t(inv_ZVZ))

  #generazione
  beta[,i]=rmvnorm(1,mubpost,sigmabpost)
  Bi=matrix(beta[,i],d*p+1+m,d,byrow=T)
  #####
  #genero i vi

  pargig3=2+lambda^2/sigma2*t(sqrt(diag(SIGMA[, , i-1])))%*%
  solve(SIGMA[, , i-1])%*%sqrt(diag(SIGMA[, , i-1]))
  eps=Y-t(Z)%*%Bi
  for(j in 1:t1){

    pargig2=1/sigma2*t(eps[j,])%*%solve(SIGMA[, , i-1])%*%eps[j,]
    v[j,i]=rgig(1,parig1,parig2,parig3)
  }

  #####
  #genero SIGMA, tramite mcmc

  #calcolo iperparametri posteriori
  Eps_tot=Y-t(Z)%*%Bi-lambda*matrix(v[,i],t1,1)%*%
  matrix(sqrt(diag(SIGMA[, , i-1])),1,d)
  Eps_omega=Eps_tot/sqrt(v[,i])
  CN=SIGMA[, , i-1]*(cn-d-1)

```

```

#genero sigma candidata
sigma_star=rinvwishart(cn,CN)

#quantita utili per calcolare alpha
Eps_tot_star=Y-t(Z)**Bi-
lambda*matrix(v[,i],t1,1)**matrix(sqrt(diag(sigma_star)),1,d)
Eps_omega_star=Eps_tot_star/sqrt(v[,i])
Cnstar=sigma_star*(cn-d-1)

#passo metropolis
alpha[i]=exp(lpost_inomega(sigma_star,
vec(t(Eps_omega_star)),Eps_omega_star,
cn,sigma2)-lpost_inomega(SIGMA[,i-1],
vec(t(Eps_omega_star)),Eps_omega_star,cn,sigma2)+
dinvwishart(SIGMA[,i-1],cn,Cnstar,log=T)-
dinvwishart(sigma_star,cn,CN,log=T))
SIGMA[,i]=SIGMA[,i-1]
if(soglie[i]<alpha[i]){
  SIGMA[,i]=sigma_star
  acc=acc+1
}
}
beta=beta[,burnin:R]
sigma=SIGMA[,burnin:R]
latent_factor=v[,burnin:R]
MM=seq(1,R-burnin,by=10)
beta=beta[,MM]
sigma=sigma[,MM]
latent_factor=latent_factor[,MM]

return(list(accepted=acc/R,beta=beta,sigma=sigma,latent_factor=
latent_factor,quantile=q,y=y,x=Z,soglia=alpha))}

```

Viene qui riportato il codice per la stima di un modello VAR quantilico con restrizioni, con gli algoritmo presentati in 4.2.3.

Codice 3: *Algoritmo stima VAR con restrizioni*

```
function(Y,X=NULL,p,q,R,burnin=5000,c,Om,om){
  #dimensioni
  d=dim(Y)[2]#t*d
  t=dim(Y)[1]
  m=dim(X)[2]#t*m
  if(is.null(X)){
    m=0
  }
  if(is.null(m)){
    m=1
  }

  #preparo le mie y e x
  Z=createZ(Y,X,p,d,t,m)

  #tolgo le prime p osservazioni
  Y=Y[(p+1):t,]
  y=vec(t(Y))

  #costanti funzione del quantile q
  lambda=(1-2*q)/(q*(1-q))
  sigma2=sqrt(2/(q*(1-q)))

  #quantit? ripetute nel gibbs
  pargig1=(2-d)/2

  M1=dim(Om)[2]
  #C=diag(rep(10,d))+matrix(5,d,d)

  #elemnti utili gibbs

  acc=0
  X=kronecker(t(Z),diag(rep(1,d)))
```

```

##adesso creo le matrici in cui porter? i miei risultati

#da qui in poi uso t1 al posto di t
t1=t-p
v=matrix(NA,t1,R)
delta=matrix(NA,M1,R)
SIGMA=array(NA,dim=c(d,d,R))
sqD=matrix(NA,d,R)

#prima iterazione
v[,1]=rexp(t1)
delta[,1]=runif(M1,-1,1)
Ri=cov2cor(cov(Y))
SIGMA[,1]=diag(sigma2)*%Ri*%diag(sigma2)
sqD[,1]=runif(d,0,5)
VARi=diag(sqD[,1])*%SIGMA[,1]*%diag(sqD[,1])
cn=t-d
cat(dim(Z))
Y=Y-matrix((kronecker(t(Z),diag(rep(1,d))))*%om),t1,d)
#mcmc
pb=txtProgressBar(0,R,initial=0)
soglie=runif(R)

for(i in 2:R){
  if(i%%50==0){setTxtProgressBar(pb,i)}
  ##genero per primo delta
  eps_tilde=Y-matrix(v[,i-1],t1,1)*% (matrix(sqD[,i-1],1,d)*lambda)
  ZV=Z*%diag(1/v[,i-1])
  ZVZ=ZV*%t(Z)
  invVARi=pseudoinverse(VARi)
  #par.posteriori delta
  sigmabpost=t(0m)*%kronecker(ZVZ,invVARi)*%0m
  sigmabpost=pseudoinverse((sigmabpost/2+t(sigmabpost)/2))
  sigmabpost=sigmabpost/2+t(sigmabpost)/2
}

```

```

mubpost=sigmapost%*(t(0m)%*vec(invVARi%*t(eps_tilde)%*t(ZV)))
delta[,i]=rmvnorm(1,mubpost,sigmapost)
XCd=X%*0m%*delta[,i]
XCd=matrix(XCd,t1,d,byrow=TRUE)
Eps=Y-XCd
#####
#genero i vi
invVARi=pseudoinverse(VARi)
pargig3=2+t(sqD[,i-1])*diag(lambda)%*invVARi
%*diag(lambda)%*sqD[,i-1]
for(j in 1:t1){
  pargig2=t(Eps[j,])%*invVARi%*Eps[j,]
  v[j,i]=rgig(1,pargig1,pargig2,pargig3)
}
#####
#genero D, tramite mh
sqDstar=sqD[,i-1]+c*runif(d,-1,1)
sqDstar=maxd(sqDstar)
VARistar=diag(sqDstar)%*SIGMA[,i-1]*diag(sqDstar)
Eps_tot_star=Eps-matrix(v[,i],t1,1)%* (matrix(sqDstar,1,d)*lambda)
Eps_omega_star=Eps_tot_star/sqrt(v[,i])
Eps_tot=Eps-matrix(v[,i],t1,1)%* (matrix(sqD[,i-1],1,d)*lambda)
Eps_omega=Eps_tot/sqrt(v[,i])
#passo metropolis
COVi=diag(sqD[,i-1])%*Ri%*diag(sqD[,i-1])
COVstar=diag(sqDstar)%*Ri%*diag(sqDstar)
alpha=exp(lpost_inomega(VARistar,vec(t(Eps_omega_star)),Eps_omega_star
, cn,COVstar)
-lpost_inomega(VARi,vec(t(Eps_omega)),Eps_omega,cn,COVi))
sqD[,i]=sqD[,i-1]
if(soglie[i]<alpha){
  sqD[,i]=sqDstar
  acc=acc+1
}
#####

```

```

#genero SIGMA
Eps_tot=Eps-matrix(v[,i],t1,1)%*(matrix(sqD[,i],1,d)*lambda)
SIGMA_post=diag(1/sigma2)%*diag(1/sqD[,i])%*t(Eps_tot)%*diag(1/v[,i
  ])%*Eps_tot%*diag(1/sqD[,i])%*diag(1/sigma2)
Ri=cov2cor(diag(1/sigma2)%*rinvwishart(t1,SIGMA_post)%*diag(1/sigma2
  ))
SIGMA[,i]=diag(sigma2)%*Ri%*diag(sigma2)
VARi=diag(sqDstar)%*SIGMA[,i]%*diag(sqDstar)
}
delta=delta[,burnin:R]
sigma=SIGMA[,burnin:R]
latent_factor=v[,burnin:R]
sqD=sqD[,burnin:R]
MM=seq(1,R-burnin,by=10)
delta=delta[,MM]
sigma=sigma[,MM]
latent_factor=latent_factor[,MM]
sqD=sqD[,MM]
return(list(accepted=acc/R,delta=delta,sigma=sigma,latent_factor=latent_
  factor,sqD=sqD,quantile=q,y=y,x=Z,soglia=alpha))}

```

Viene qui riportato il codice per la stima di un modello VAR con priori *spike-and-slab* come quello presentato in 4.2.4.

Codice 4: *Algoritmo stima spike-and-slab*

```

function(Y,X=NULL,p,q,R=20000,burnin=10000,c=1,a_u1=2.1,b_u1=4,a_seps=2.1,
  b_seps=4,u0=0.001,a_pi0=2,b_pi0=2){
#dimensioni
d=dim(Y)[2]#t*d
t=dim(Y)[1]
m=dim(X)[2]#t*m
if(is.null(X)){
  m=0
}
#preparo le mie y e x
Z=createZ(Y,X,p,d,t,m)
Z1=Z[1,]

```

```
Z=Z[-1,]

#tolgo le prime p osservazioni
Y=Y[(p+1):t,]
y=vec(t(Y))
#costanti funzione del quantile q
lambda=(1-2*q)/(q*(1-q))
sigma2=sqrt(2/(q*(1-q)))

#quantita ripetute nel gibbs
pargig1=(2-d)/2

#C=diag(rep(10,d))+matrix(5,d,d)
#elemnti utili gibbs
acc=0
M=d*d*p+d*m
##adesso creo le matrici in cui portero i miei risultati

#da qui in poi uso t1 al posto di t
t1=t-p
v=matrix(NA,t1,R)
mu=matrix(NA,d,R)
beta=matrix(NA,M,R)
SIGMA=array(NA,dim=c(d,d,R))
sqD=matrix(NA,d,R)
sigeps=matrix(NA,d*p+m,R)
pi0=matrix(NA,d,R)
gamma=array(NA,c(d,d*p+m,R))
u1=matrix(NA,d,R)

#prima iterazione
v[,1]=rexp(t1)
u1[,1]=rinvgamma(d,a_u1,b_u1)
sigeps[,1]=rinvgamma(d*p+m,a_seps,b_seps)
pi0[,1]=rbeta(d,a_pi0,b_pi0)
```

```

gamma[, , 1]=t(rbinom(d*p+m,1,pi0[, 1]))

#preparo sigmab
sigmab_mat=sigeps[, 1]*(gamma[, , 1]+(1-gamma[, , 1])*u0)*u1[, 1]
cat(dim(Z))
sigmabvec=vec(sigmab_mat)
sigmab=diag(sigmabvec)
sigmabinv=solve(sigmab)
beta[, 1]=rmvnorm(1,rep(0,M),sigmab)
Bi=matrix(beta[, 1],d*p+m,d,byrow=TRUE)
mu[, 1]=runif(d)

#preparo la varianza dei dati
Ri=cov2cor(cov(Y))
SIGMA[, , 1]=diag(sigma2)%*%Ri%*%diag(sigma2)
sqD[, 1]=sqrt(diag(diag(1/sigma2)%*%diag(runif(d,0,5))%*%diag(1/sigma2)))
VARi=diag(sqD[, 1])%*%SIGMA[, , 1]%*%diag(sqD[, 1])
cn=t1
invVARi=solve(VARi)
#mcmc
pb=txtProgressBar(0,R,initial=0)
soglie=runif(R)

for(i in 2:R){
  if(i%%50==0){setTxtProgressBar(pb,i)}
  ##genero per primo beta
  eps_tilde=Y-matrix(v[, i-1],t1,1)%*% matrix(sqD[, i-1],1,d)*lambda-
    matrix(mu[, i-1],t1,d,byrow = T)
  ZV=Z%*%diag(1/v[, i-1])
  ZVZ=ZV%*%t(Z)
  #par.posteriori beta
  sigmabpost=pseudoinverse(sigmabinv+kronecker(ZVZ,invVARi))
  sigmabpost=(sigmabpost/2+t(sigmabpost)/2)
  mubpost=sigmabpost%*%(vec(invVARi%*%t(eps_tilde)%*%t(Z)))
  #generazione
  beta[, i]=rmvnorm(1,as.vector(mubpost),sigmabpost)
}

```

```

Bi=matrix(beta[,i],d*p+m,d,byrow=T)

#Genero da sigeps
for(j in 1:(m+d*p)){
  post_sigeps=1/2*sum(Bi[j,]^2/((gamma[,j,i-1]+(1-gamma[,j,i-1])*u0)*
    u1[,i-1]))+b_seps
  sigeps[j,i]=rinvgamma(1,a_seps+(d)/2,post_sigeps)}

#genero gamma
for(l in 1:d){
  for(j in 1:(m+d*p)){
    p1=exp(-Bi[j,l]^2/(2*sigeps[j,i]*u1[l,i-1]))*pi0[l,i-1]/sqrt(
      sigeps[j,i]*u1[l,i-1])
    p0=exp(-Bi[j,l]^2/(2*sigeps[j,i]*u1[l,i-1]*u0))*(1-pi0[l,i-1])/
      sqrt(sigeps[j,i]*u1[l,i-1]*u0)
    p1_post=p1/(p1+p0)
    p1_post[is.na(p1_post)]=0.5
    gamma[l,j,i]=rbinom(1,1,p1_post)}}

totgammai=apply(gamma[, , i], 1, sum)

#genero pi0
pi0[,i]=rbeta(1,a_pi0+totgammai,m+p*d+b_pi0-totgammai)

#genero u1
for(l in 1:d){
  quant2=1/2*apply((Bi[,l]^2/((gamma[l, , i-1]+(1-gamma[l, , i-1])*u0)*t(
    sigeps[,i]))),1,sum)+b_u1#metti a posto posteriori
  u1[l,i]=rinvgamma(1,a_u1+M/2,quant2)}
sigmab_mat=sigeps[,i]*(gamma[, , i]+(1-gamma[, , i])*u0)*u1[,i]
sigmabvec=vec(sigmab_mat)
sigmab=diag(sigmabvec)
sigmabinv=solve(sigmab)

#genero mu
eps_tilde=Y-matrix(v[,i-1],t1,1)%*% matrix(sqD[,i-1],1,d)*lambda-t(Z)%

```

```

    *%Bi

#par.posteriori mu
sigmamupost=VARi/sum(1/v[,i-1])
sigmamupost=(sigmamupost/2+t(sigmamupost)/2)
mumupost=t(eps_tilde)%*(1/v[,i-1])/sum(1/v[,i-1])

#generazione
mu[,i]=rmvnorm(1,mumupost,sigmamupost)
#####
#genero i vi
invVARi=pseudoinverse(VARi)
pargig3=2+t(lambda)%*diag(sqD[,i-1])%*invVARi%*diag(sqD[,i-1])%*
    lambda
eps=Y-t(Z)%*Bi-matrix(mu[,i],t1,d,byrow = T)
for(j in 1:t1){
    pargig2=t(eps[j,])%*invVARi%*eps[j,]
    v[j,i]=rgig(1,pargig1,pargig2,pargig3)
}

#####
#genero D, tramite mh
sqDstar=sqD[,i-1]+c*runif(d,-1,1)
sqDstar=maxd(sqDstar)
VARistar=diag(sqDstar)%*SIGMA[,i-1]%*diag(sqDstar)
Eps_tot_star=Y-t(Z)%*Bi-matrix(v[,i],t1,1)%* (matrix(sqDstar,1,d)*
    lambda)-matrix(mu[,i],t1,d,byrow = T)
Eps_omega_star=Eps_tot_star/sqrt(v[,i])
Eps_tot=Y-t(Z)%*Bi-matrix(v[,i],t1,1)%* (matrix(sqD[,i-1],1,d)*
    lambda)-matrix(mu[,i],t1,d,byrow = T)
Eps_omega=Eps_tot/sqrt(v[,i])
#passo metropolis

COVi=diag(sqD[,i-1])%*Ri%*diag(sqD[,i-1])
COVstar=diag(sqDstar)%*Ri%*diag(sqDstar)
alpha=exp(lpost_inomega(VARistar,vec(t(Eps_omega_star)),Eps_omega_star

```

```

    ,cn,COVstar)
        -lpost_inomega(VARi,vec(t(Eps_omega)),Eps_omega,cn,COVi))
sqD[,i]=sqD[,i-1]
if(soglie[i]<alpha){
  sqD[,i]=sqDstar
  acc=acc+1
}
#####
#genero SIGMA
Eps_tot=Y-t(Z)%*%Bi-matrix(v[,i],t1,1)%*% (matrix(sqD[,i],1,d)*lambda)
  -matrix(mu[,i],t1,d,byrow = T)

SIGMA_post=diag(1/sqD[,i])%*%t(Eps_tot)%*%diag(1/v[,i])%*%Eps_tot%*%
  diag(1/sqD[,i])
SIGMA_post=nearPD(SIGMA_post)$mat
Rinv=rinvwishart(t1,SIGMA_post)
Ri=diag(1/sigma2)%*%Rinv%*%diag(1/sigma2)
Ri=diag(1/sqrt(diag(Ri)))%*%Ri%*%diag(1/sqrt(diag(Ri)))
SIGMA[,i]=diag(sigma2)%*%Ri%*%diag(sigma2)
VARi=diag(sqDstar)%*%SIGMA[,i]%*%diag(sqDstar)

}
beta=beta[,burnin:R]
sigma=SIGMA[,burnin:R]
latent_factor=v[,burnin:R]
sqD=sqD[,burnin:R]
mu=mu[,burnin:R]
gamma=gamma[,burnin:R]
prob=pi0[,burnin:R]
u1=u1[,burnin:R]
sigeps=sigeps[,burnin:R]
MM=seq(1,R-burnin,by=10)
beta=beta[,MM]
sigma=sigma[,MM]
latent_factor=latent_factor[,MM]

```

```

sqD=sqD[,MM]
mu=mu[,MM]
gamma=gamma[, ,MM]
prob=pi0[,MM]
u1=u1[,MM]
sigeps=sigeps[,MM]
return(list(accepted=acc/R,beta=beta,sigma=sigma,mu=mu,latent_factor=
  latent_factor,sqD=sqD,gamma=gamma,prob=pi0,u1=u1,sigeps=sigeps,
  quantile=q,y=y,x=Z,soglia=alpha))}

```

Vengono infine riportati i codici per lo svolgimento delle analisi strutturali quantiliche presentate in 3.5.1.

Codice 5: *Analisi strutturali quantiliche*

```

QIRF=function(rit,B_ar,D,Omega,vmed,d,p,j){
  ej=rep(0,d)
  ej[j]=1
  J=cbind(diag(c(rep(1,d))),matrix(0,d,d*(p-1)))
  A=diag(rep(1,d*(p-1)))
  A=cbind(A,matrix(0,d*(p-1),d))
  A=rbind(B_ar,A)
  IRFi=matrix(NA,d,rit)
  omega12=sqrt(vmed)*D%*%t(chol(Omega))
  for(i in 1:rit){
    IRFi[,i]=J%*%(A^i)%*%t(J)%*%(omega12)%*%ej
  }
  IRFi
}

```

```

GQIRF=function(rit,B_ar,D,Omega,vmed,d,p,j){
  ej=rep(0,d)
  ej[j]=1
  J=cbind(diag(c(rep(1,d))),matrix(0,d,d*(p-1)))
  A=diag(rep(1,d*(p-1)))
  A=cbind(A,matrix(0,d*(p-1),d))
  A=rbind(B_ar,A)
}

```

```

Djj=D[j,j]*chol(Omega)[j,j]
ej=ej/Djj
IRFi=matrix(NA,d,rit)
  for(i in 1:rit){
    IRFi[,i]=J**(A^i)**t(J)**D**(sqrt(vmed)*Omega)**D**ej
  }
  IRFi
}

QFEVD=function(rit,B_ar,D,Omega,vmed,d,p,j){
  ej=rep(0,d)
  ej[j]=1
  J=cbind(diag(c(rep(1,d))),matrix(0,d,d*(p-1)))
  A=diag(rep(1,d*(p-1)))
  A=cbind(A,matrix(0,d*(p-1),d))
  A=rbind(B_ar,A)
  IRFi=matrix(NA,d,rit)
  omega12=sqrt(vmed)*D**t(chol(Omega))
  for(i in 1:rit){
    for(k in 1:d){
      ek=rep(0,d)
      ek[k]=1
      IRFi[k,i]=(ej**J**(A^i)**t(J)**(omega12)**ek)^2
    }
  }
  apply(IRFi,1,sum)/sum(IRFi)
}

GQFEVD=function(rit,B_ar,D,Omega,vmed,d,p,j){
  ej=rep(0,d)
  ej[j]=1
  J=cbind(diag(c(rep(1,d))),matrix(0,d,d*(p-1)))
  A=diag(rep(1,d*(p-1)))
  A=cbind(A,matrix(0,d*(p-1),d))
  A=rbind(B_ar,A)
  IRFi=matrix(NA,d,rit)
  omega12=sqrt(vmed)*D**t(chol(Omega))

```

```
for(i in 1:rit){
  for(k in 1:d){
    ek=rep(0,d)
    ek[k]=1
    Dkk=D[k,k]*chol(Omega)[k,k]
    ek=ek/Dkk
    IRFi[k,i]=(ej**J**(A^i)**t(J)**(omega12)**ek)^2
  }
}
apply(IRFi,1,sum)/sum(IRFi)}
```