

Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Magistrale in
Scienze Statistiche



**Studio e Applicazione di Modelli Autoregressivi per
Serie Temporali a Valori Matriciali**

Relatore Prof. Massimiliano Caporin
Dipartimento di Scienze Statistiche

Laureando: Andrea Fiumanò
Matricola N 2026573

Anno Accademico 2023/2024

Indice

Elenco delle figure	iii
Abstract	v
Introduzione	vii
1 Fondamenti del Modello Autoregressivo per Serie Temporali a Valori Matriciali . .	1
2 Stima parametri modello MAR(1)	5
3 Caso di studio	9
Conclusioni	25
Bibliografia	27

Elenco delle figure

3.1	Serie temporali mensili delle differenze prime logaritmiche.	9
3.2	Errore di stima ridge al variare di $(\lambda_{A_R}, \lambda_{B_R})$	12
3.3	Errore di stima ridge al variare di $(\lambda_{A_R}, \lambda_{B_R})$ e valore ottimale $(\lambda_{A_R}, \lambda_{B_R})$	13
3.4	Errore di stima ridge al variare di λ_{B_R} con valore ottimale di λ_{A_R}	13
3.5	Errore di stima ridge al variare di λ_{A_R} con valore ottimale di λ_{B_R}	14
3.6	Coefficienti matrice \hat{B}_R al variare di $(\lambda_{A_R}, \lambda_{B_R})$	15
3.7	Coefficienti matrice \hat{A}_R al variare di $(\lambda_{A_R}, \lambda_{B_R})$	15
3.8	Coefficienti matrice \hat{B}_R al variare di λ_{B_R} con valore ottimale di λ_{A_R}	16
3.9	Coefficienti matrice \hat{B}_R al variare di λ_{A_R} con valore ottimale di λ_{B_R}	16
3.10	Coefficienti matrice \hat{A}_R al variare di λ_{B_R} con valore ottimale di λ_{A_R}	16
3.11	Coefficienti matrice \hat{A}_R al variare di λ_{A_R} con valore ottimale di λ_{B_R}	16
3.12	Errore di stima ridge al variare di λ_{B_R}	17
3.13	Coefficienti matrice \hat{B}_R al variare di λ_{B_R}	17
3.14	Errore di stima lasso al variare di λ_{B_L}	18
3.15	Coefficienti matrice \hat{B}_L al variare di λ_{B_L}	19
3.16	Funzione di Autocorrelazione degli errori con stimatori \hat{A}_{LSE} e \hat{B}_{LSE}	22
3.17	Funzione di Autocorrelazione degli errori con stimatori \hat{A}_R e \hat{B}_R	22
3.18	Funzione di Autocorrelazione degli errori con stimatori \hat{A}_L e \hat{B}_L	23
3.19	Grafici degli errori con stimatori \hat{A}_{LSE} e \hat{B}_{LSE}	23
3.20	Grafici degli errori degli errori con stimatori \hat{A}_R e \hat{B}_R	23
3.21	Grafici degli errori degli errori con stimatori \hat{A}_L e \hat{B}_L	24

Abstract

La presente tesi si concentra sull'analisi e sull'applicazione pratica dei modelli autoregressivi per serie temporali con valori matriciali, fornendo un quadro teorico e metodologico per affrontare questa complessa sfida analitica. Attraverso una revisione dettagliata della letteratura, la ricerca conferma la validità del modello autoregressivo in questo contesto. Un caso di studio riguardante alcune serie storiche economiche relative a diverse nazioni dell'area euro viene esaminato per illustrare la rilevanza e l'efficacia del modello proposto. In sintesi, questa tesi rappresenta un contributo significativo alla comprensione e all'applicazione dei modelli autoregressivi nell'ambito di serie temporali con valori matriciali, aprendo nuove prospettive per l'analisi di dati complessi e multidimensionali. La sua implementazione pratica fornisce un valido strumento per l'analisi e la previsione di serie temporali caratterizzate da questa particolare struttura dati.

Introduzione

Le serie temporali multivariate rappresentano un ambito molto comune nell'analisi delle serie storiche, e di recente, si è manifestato un interesse crescente nel modellare serie temporali caratterizzate da una elevata dimensionalità. Questi studi si suddividono in diverse categorie, tra cui ad esempio l'applicazione di modelli autoregressivi vettoriali con tecniche di regolarizzazione, o l'utilizzo di modelli statistici o dinamici a fattori. Nella maggior parte dei casi, le varie osservazioni ad ogni istante temporale vengono trattate come un vettore. Tuttavia, nonostante la convenzionale rappresentazione vettoriale, l'interconnessione tra le serie temporali spesso presenta una struttura più complessa. Quando le serie temporali sono organizzate secondo le intersezioni di due classificazioni, si presentano naturalmente sotto forma di matrici. Sebbene l'analisi delle serie temporali univariate tratti singolarmente ciascuna variabile e l'analisi delle serie temporali panel consideri le variabili una riga o una colonna alla volta, è ragionevole presumere che, ad esempio, variabili economiche simili provenienti da paesi differenti possiedano una forte relazione tra di loro. Allo stesso tempo, è plausibile che gli indicatori economici di uno stesso paese presentino movimenti sincronizzati. Pertanto, emerge una struttura robusta nelle relazioni tra le serie temporali. Concatenando le matrici in vettori, si rischia di perdere tale struttura sottostante, con conseguenti impatti significativi sulla complessità del modello e sulle interpretazioni.

È possibile adattare un particolare modello autoregressivo concepito per la modellazione delle serie temporali con valori matriciali, utilizzando un approccio bilineare. Questo modello è stato studiato ed implementato da Chen, Rong and Xiao, Han and Yang, Dan nel loro articolo [2].

Nello specifico, la stima della media condizionale dell'osservazione matriciale al tempo t viene ottenuta tramite la moltiplicazione della matrice osservata al tempo precedente $t - 1$, sia a sinistra che a destra, per due matrici di coefficienti autoregressivi.

Sia X_t la matrice $m \times n$ osservata al tempo t , il modello autoregressivo di ordine 1 per serie temporali a valori matriciali, indicato con l'acronimo MAR(1) assume la forma:

$$X_t = AX_{t-1}B' + E_t \quad (1)$$

Il modello può essere esteso coinvolgendo le precedenti p matrici osservate per formare un modello autoregressivo di ordine p . Nonostante ciò, ci si concentrerà solo sullo studio del modello MAR di ordine 1.

Rispetto ai tradizionali modelli autoregressivi vettoriali, questo approccio presenta alcuni importanti vantaggi:

- Sfrutta appieno la struttura originale della matrice, in modo che il modello sia naturalmente interpretabile.
- Le matrici dei coefficienti A e B hanno interpretazioni corrispondenti.
- Si riduce significativamente il numero di parametri rispetto a un modello autoregressivo vettoriale diretto ($m^2 + n^2$ rispetto a m^2n^2).

Quando le osservazioni delle matrici stesse hanno dimensioni elevate, sarebbe opportuno introdurre vincoli aggiuntivi per ottenere una riduzione delle dimensioni più significativa. Esistono diversi possibili approcci. In questa tesi verranno utilizzate alcune funzioni di penalizzazione in fase di stima dei parametri per ovviare a tale inconveniente.

La presente tesi si propone quindi di esaminare il modello autoregressivo per serie temporali a valori matriciali di ordine 1, in particolare prendendo spunto dal testo di riferimento e riportando le parti più importanti.

Il primo capitolo della tesi sarà dedicato all'introduzione del modello autoregressivo per serie temporali a valori matriciali, concentrandosi sul modello MAR(1). Saranno esaminati in dettaglio i concetti chiave del modello, comprese le interpretazioni delle matrici dei coefficienti e le proprietà probabilistiche del modello.

Nel secondo capitolo verrà presentato lo stimatore dei minimi quadrati implementato dagli autori per il modello MAR, accompagnato da due ulteriori stimatori che verranno calcolati ponendo un termine di penalizzazione all'interno della funzione dei minimi quadrati.

Il terzo capitolo sarà dedicato alla trattazione di un caso di studio specifico, focalizzato su dati relativi ad indicatori economici relativi a diverse nazioni europee. Attraverso l'impiego del modello MAR(1), si procederà con un'analisi dei dati.

Capitolo 1

Fondamenti del Modello Autoregressivo per Serie Temporali a Valori Matriciali

Nel presente capitolo, si pongono le fondamenta per la comprensione del Modello Autoregressivo per Serie Temporali a Valori Matriciali di ordine 1, che rappresenta il fulcro della tesi. Attraverso una definizione formale del modello, si analizza sia la struttura della matrice degli errori che quella della matrice di covarianza ad essa associata. Viene approfondita l'interpretazione del modello, con particolare riferimento alle matrici dei coefficienti. Infine vengono esposte le proprietà probabilistiche del modello. Questo approccio iniziale fornirà una solida base concettuale per analizzare in dettaglio le serie temporali che saranno oggetto di studio nel caso di ricerca.

Definizione: Si consideri una serie temporale di lunghezza T . Sia X_t la matrice $m \times n$ osservata al tempo t , per $t = 1, \dots, T$, il modello autoregressivo di ordine 1 per serie temporali a valori matriciali, indicato con l'acronimo MAR(1), è definito dalla seguente equazione:

$$X_t = AX_{t-1}B' + E_t, \quad (1.1)$$

dove $A = (a_{ij})$ e $B = (b_{ij})$ sono matrici di coefficienti autoregressivi ($m \times m$) e ($n \times n$), rispettivamente, e $E_t = (\epsilon_{ij,t})$ è una matrice white noise $m \times n$.

Questo modello può essere reinterpretato come un caso particolare di un modello vettoriale autoregressivo di ordine 1, evidenziando la struttura dei coefficienti tramite il prodotto di Kronecker tra matrici.

$$\text{vec}(X_t) = (B \otimes A)\text{vec}(X_{t-1}) + \text{vec}(E_t), \quad (1.2)$$

Tuttavia, è importante sottolineare che il modello MAR(1) richiede un numero inferiore di coefficienti rispetto a un VAR(1) generico. Mentre il modello MAR(1) richiede $m^2 + n^2$ coefficienti, dove m e n rappresentano le dimensioni delle variabili coinvolte, un VAR(1) non vincolato richiede un numero di coefficienti pari a m^2n^2 . Questa differenza diventa particolarmente rilevante in contesti in cui le dimensioni delle variabili sono considerevoli.

Un aspetto critico da considerare riguarda invece l'identificabilità delle matrici di coefficienti nel modello MAR(1). È noto che il modello rimane invariato se entrambe le matrici di coefficienti, A e B , vengono moltiplicate per lo stesso valore costante non nullo. Per risolvere questo problema e garantire una soluzione unica, si adotta la pratica di normalizzare la matrice A in modo che la sua norma di Frobenius sia pari a uno. Allo stesso tempo, la soluzione unica è sempre garantita per il prodotto di Kronecker ($B \otimes A$). Queste considerazioni sono cruciali per garantire l'affidabilità e l'interpretazione corretta del modello MAR(1) nelle analisi.

In riferimento alla matrice del termine di errore E_t , si presuppone che la sequenza di matrici di errore $\{E_t\}_{t=1,\dots,T}$ segua un processo White Noise matriciale, caratterizzato dall'assenza di correlazione in istanti temporali differenti, ossia

$$\text{Corr}(E_t, E_s) = 0 \quad \forall t \neq s.$$

Tuttavia, è importante notare che E_t può comunque presentare correlazioni simultanee tra le sue voci, il che rende la comprensione della struttura della covarianza una sfida ardua. Per semplificare, ci si avvale della seguente rappresentazione per la matrice di covarianza Σ :

$$\Sigma = \text{Cov}(\text{vec}(E_t)), \quad (1.3)$$

la quale ha dimensioni $(mn) \times (mn)$.

Nel caso più semplice, è possibile ipotizzare indipendenza tra le voci di E_t , il che conduce a una matrice di covarianza diagonale. Tuttavia, per esplorare scenari più realistici, sono ammesse correlazioni arbitrarie tra le voci.

Il modello MAR(1) non è di facile interpretazione e richiede un'analisi accurata per comprendere appieno le implicazioni delle matrici dei coefficienti. Si osserva che all'interno dell'equazione (1.1), la matrice A a sinistra riflette le interazioni riga per riga, mentre la matrice B a destra introduce una dipendenza colonna per colonna. Di conseguenza, la media condizionale nel modello MAR(1) combina sia le interazioni riga per riga che colonna per colonna. L'analisi di casi particolari aiuta a rendere più chiara la comprensione di come le matrici dei coefficienti A e B riflettano le strutture delle righe e delle colonne.

Isolando l'effetto della forma bilineare nel modello MAR(1) attraverso l'assunzione, ad esempio, di $A = I$, ovvero la matrice identità. Ciò conduce a un'equazione semplificata del modello, che si riduce a:

$$X_t = X_{t-1}B' + E_t. \quad (1.4)$$

Questa relazione implica che, al tempo t , la media condizionale di un elemento di riga della matrice dei dati X_t per una data colonna è espressa come una combinazione lineare dello stesso elemento relativo a tutte le colonne di X_{t-1} . Tale combinazione lineare è costante per ogni riga. Di conseguenza, questo caso particolare del modello e la matrice B catturano le interazioni tra le colonne. Tuttavia, tali interazioni sono circoscritte all'interno di ciascuna riga, mancando di interazioni tra righe diverse. Inoltre, è possibile esprimere ogni riga di X_t come

$$X_{t,i} = X_{t-1,i}B' + E_{t,i}, \quad i = 1, \dots, m. \quad (1.5)$$

In questo caso ogni riga di X_t segue lo stesso modello VAR(1) di dimensione n . Inoltre lo stesso modello VAR(1) è valido per ogni riga della matrice dei dati.

D'altra parte, nel caso in cui si ponga la matrice B uguale alla matrice identità I , il modello MAR(1) (1.1) si riduce a:

$$X_t = AX_{t-1} + E_t. \quad (1.6)$$

Nello scenario relativo all'equazione (1.6), la matrice A descrive le interazioni tra le righe. In tal caso, come è logico aspettarsi, non vi sono interazioni tra colonne di X_t . La media condizionale di un elemento colonna della matrice dei dati per una data riga è espressa come una combinazione lineare dello stesso elemento relativo a tutte le righe di X_{t-1} .

In secondo luogo è possibile esprimere ogni colonna di X_t come,

$$X_{t,j} = AX_{t-1,j} + E_{t,j}, \quad j = 1, \dots, n. \quad (1.7)$$

In questo caso ogni colonna di X_t segue lo stesso modello VAR(1) m -dimensionale. Inoltre lo stesso modello VAR(1) è valido per ogni altra colonna.

Tuttavia, come era logico aspettarsi, ponendo questo vincolo sulla matrice B , i vari modelli VAR(1) enfatizzano solamente le interazione tra le colonne della matrice dei dati.

Logicamente i due modelli proposti dalle equazioni (1.4) e (1.6) risultano essere troppo restrittivi. Nonostante ciò, l'utilizzo dei vincoli sulle matrici A e B si rivela essere un'ottima idea interpretare il modello attraverso una struttura gerarchica ben definita. In particolare, si mostra che il modello condivide un'interpretazione simile alla regressione autoregressiva gerarchica.

Supponendo che,

$$Y_{t-1} = X_{t-1}B', \quad (1.8)$$

Y_{t-1} non è altro che la media condizionale di X_t se $A = I$; e ogni colonna di Y_{t-1} è basata sulla combinazione lineare di tutte le colonne di X_{t-1} senza interazioni tra le righe. Dopodichè, si considera

$$Z_{t-1} = AX_{t-1}, \quad (1.9)$$

Z_{t-1} non è altro che la media condizionale di X_t se $B = I$; e ogni riga di Z_{t-1} è basata sulla combinazione lineare di tutte le righe di X_{t-1} senza interazioni tra le colonne.

Mettendo assieme le equazioni (1.8) e (1.9) si ottiene

$$X_t = Z_{t-1} + E_t = AY_{t-1} = AX_{t-1}B' + E_t. \quad (1.10)$$

Ogni entrata $x_{t,ij}$ di X_t per $i = 1, \dots, m$ e $j = 1, \dots, n$ è data da

$$x_{t,ij} = \sum_{k_1=1}^m \sum_{k_2=1}^n a_{ik_1} x_{t-1,k_1k_2} b_{jk_2} e_{t,ij}. \quad (1.11)$$

Ogni entrata $x_{t,ij}$ è controllata solo dalla i -esima riga di A e dalla j -esima colonna di B' . In riferimento alla Figura ??, la i -esima riga di A può essere considerata come il coefficiente corrispondente all' i -esimo indice e la j -esima colonna di B' come il coefficiente corrispondente al j -esimo paese. Un'interpretazione dei singoli coefficienti più approfondita verrà vista nel capitolo 3, relativo al caso di studio.

Gli autori hanno identificato la seguente condizione necessaria e sufficiente affinché un modello MAR(1) risulti stazionario e causale.

Proposizione 1.0.1. *Dato il modello MAR(1) (1.1), se $\rho(A) \cdot \rho(B) < 1$, allora il modello ammette una soluzione stazionaria e causale; dove $\rho(A)$, ossia il raggio spettrale della matrice A , è definito come il massimo valore assoluto degli autovalori di A .*

$$\rho(A) = \max(|\lambda_1|, \dots, |\lambda_n|).$$

Quando è soddisfatta la condizione nella proposizione (1.0.1) il modello MAR(1) ammette la seguente rappresentazione causale:

$$vec(X_t) = \sum_{k=0}^{\infty} (B^k \otimes A^k) vec(E_{t-k}), \quad (1.12)$$

attraverso la quale viene espressa la funzione di autocovarianza del modello

$$\Gamma_k = Cov(vec(X_t), vec(X_{t-k})) = \sum_{l=0}^{\infty} (B^{k+l} \otimes A^{k+l}) \Sigma (B^l \otimes A^l)', \quad k \geq 0 \quad (1.13)$$

La condizione $\rho(A) \cdot \rho(B) < 1$ garantisce che la serie infinita di matrici sia assolutamente sommabile.

Per quanto riguarda l'identificabilità delle matrici A e B , specificando il seguente spazio dei parametri relativo alle due matrici, è garantita l'unicità di A e B a meno del segno.

$$\{(A, B) : \|A\|_F = 1, \|B\|_F > 0\}.$$

Capitolo 2

Stima parametri modello MAR(1)

Per la stima delle matrici dei coefficienti A e B si fa riferimento all'algoritmo dei minimi quadrati iterati, implementato dagli autori. Si assume che la sequenza di matrici relative al termine di errore E_t in (1.1) ossia $\{E_t\}_{t=1, \dots, T}$, siano indipendenti ed identicamente distribuite secondo una distribuzione normale con media zero e varianza costante:

$$E_t \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma^2), \quad t = 1, \dots, T$$

Gli stimatori della massimo verosimiglianza di A e B , denotati con \hat{A}_{LSE} e \hat{B}_{LSE} , sono la soluzione del problema dei minimi quadrati, indicato con LSE.

$$(\hat{A}_{LSE}, \hat{B}_{LSE}) = \underset{A, B}{\operatorname{arg\,min}} \left\{ \sum_{t=2}^T \|X_t - AX_{t-1}B'\|_F^2 \right\} \quad (2.1)$$

Per risolvere il problema (2.1), innanzitutto vengono calcolate le derivate parziali, rispetto ad A e B , dell'equazione $\sum_{t=2}^T \|X_t - AX_{t-1}B'\|_F^2$ in modo tale da ottenere la condizione sul gradiente per la stima ai minimi quadrati:

$$\begin{aligned} \sum_{t=2}^T AX_{t-1}B'BX'_{t-1} - \sum_{t=2}^T X_tBX_{t-1} &= 0, \\ \sum_{t=2}^T BX'_{t-1}A'AX'_{t-1} - \sum_{t=2}^T X'_tAX_{t-1} &= 0. \end{aligned} \quad (2.2)$$

Gli autori nel loro articolo [2] garantiscono e dimostrano l'esistenza e l'unicità del minimo globale per le equazioni (2.2).

Per risolvere (2.1) è stata proposta una procedura di aggiornamento iterativa, aggiornando sequenzialmente una matrice alla volta tenendo fissa l'altra, iniziando con un valore iniziale per ciascuna. Attraverso la (2.2) l'iterazione di aggiornamento di \hat{B}_{LSE} data \hat{A}_{LSE} è

$$\hat{B}_{LSE} = \left(\sum_{t=2}^T X'_t \hat{A}_{LSE} X_{t-1} \right) \left(\sum_{t=2}^T X'_{t-1} \hat{A}'_{LSE} \hat{A}_{LSE} X_{t-1} \right)^{-1}, \quad (2.3)$$

mentre l'iterazione di aggiornamento di \hat{A}_{LSE} data \hat{B}_{LSE} è

$$\hat{A}_{LSE} = \left(\sum_{t=2}^T X_t \hat{B}_{LSE} X'_{t-1} \right) \left(\sum_{t=2}^T X_{t-1} \hat{B}'_{LSE} \hat{B}_{LSE} X'_{t-1} \right)^{-1}. \quad (2.4)$$

L'algoritmo iterativo termina quando si arriva a convergenza, ossia quando ciascun elemento di \hat{A}_{LSE} e \hat{B}_{LSE} al passo j -esimo risulta essere uguale al suo corrispondente al passo precedente. Sono necessarie due matrici di partenza: in questa tesi ed in particolare nel caso di studio verranno utilizzate le matrici identiche I di dimensione $(m \times m)$ e $(n \times n)$. Inoltre, ogni volta che si aggiorna la matrice \hat{A}_{LSE} , essa viene normalizzata in modo tale che la sua norma di Frobenius sia uguale ad 1.

Partendo dal problema dei minimi quadrati in equazione (2.1), è possibile introdurre un termine di penalizzazione legato alle matrici dei coefficienti. Ad esempio, facendo riferimento alla regressione con termine di penalizzazione di tipo ridge è possibile formulare il problema dei minimi quadrati penalizzati come segue:

$$(\hat{A}_R, \hat{B}_R) = \arg \min_{A, B} \left\{ \sum_{t=2}^T \|X_t - AX_{t-1}B'\|_F^2 + \lambda_{A_R} \|A\|_F^2 + \lambda_{B_R} \|B\|_F^2 \right\}, \quad (2.5)$$

dove $\lambda_{A_R} > 0$ e $\lambda_{B_R} > 0$.

\hat{A}_R e \hat{B}_R indicano rispettivamente le stime di A e B ottenuta mediante il metodo dei minimi quadrati con penalizzazione di tipo ridge. λ_{A_R} e λ_{B_R} sono detti parametri di penalizzazione, al variare di essi varia il valore dei coefficienti stimati delle matrici A e B . All'aumentare del valore di λ_{A_R} gli elementi di \hat{A}_R subiscono una contrazione verso lo zero. Ipotizzando $\hat{A}_R \rightarrow \infty$, la matrice \hat{A}_R avrebbe ogni elemento tendente a zero. In generale, vengono maggiormente contratti verso lo zero i coefficienti relativi alla parte dei dati che spiega meno variabilità rispetto a quelli utili a spiegare più variabilità. D'altro canto ponendo $\hat{A}_R = 0$, si rimuove la penalizzazione sulla matrice A , quindi la stima dei coefficienti non subirà alcuna contrazione verso lo zero. Ovviamente lo stesso discorso risulta altrettanto valido per λ_{B_R} in relazione alla matrice \hat{B}_R . Per risolvere l'equazione (2.9), come nel caso della (2.1), si calcolano le sue derivate parziali rispetto ad A e B e si pongono a zero:

$$\begin{aligned} \sum_{t=2}^T AX_{t-1}B'BX'_{t-1} - \sum_{t=2}^T X_tBX_{t-1} + \lambda_{A_R}A &= 0, \\ \sum_{t=2}^T BX'_{t-1}A'AX'_{t-1} - \sum_{t=2}^T X'_tAX_{t-1} + \lambda_{B_R}B &= 0. \end{aligned} \quad (2.6)$$

Siccome è possibile esplicitare A e B nell'equazione (2.10), si utilizza anche qui una procedura di aggiornamento iterativa per trovare \hat{A}_R e \hat{B}_R , utilizzando le matrici identità I_m e I_n come valori iniziali.

L'iterazione di aggiornamento di \hat{B}_R data \hat{A}_R è la seguente:

$$\hat{B}_R = \left(\sum_{t=2}^T X'_t \hat{A}_R X_{t-1} \right) \left(\sum_{t=2}^T X'_{t-1} \hat{A}'_R \hat{A}_R X_{t-1} + \lambda_{B_R} I_n \right)^{-1}, \quad (2.7)$$

mentre l'iterazione di aggiornamento di \hat{A}_R data \hat{B}_R è

$$\hat{A}_R = \left(\sum_{t=2}^T X_t \hat{B}_R X'_{t-1} \right) \left(\sum_{t=2}^T X_{t-1} \hat{B}'_R \hat{B}_R X'_{t-1} + \lambda_{A_R} I_m \right)^{-1}. \quad (2.8)$$

L'algoritmo iterativo anche in questo termina quando si arriva a convergenza e ad ogni passo si normalizza \hat{A}_R in modo tale che la sua norma di Frobenius sia uguale ad 1. Non esiste un valore predefinito da assegnare a λ_{A_R} e λ_{B_R} , tuttavia è possibile trovare un valore ottimale per entrambi attraverso il metodo di stima e verifica: una prima parte dei dati è utilizzata per implementare il modello con vari valori per i parametri di penalizzazione; successivamente si seleziona la coppia di valori per la quale si ha il minimo errore di previsione; dopodichè si adatta il modello alla seconda parte dei dati con la coppia $(\lambda_{A_R}, \lambda_{B_R})$ trovata in precedenza.

Si può introdurre anche un'altra funzione di penalizzazione legata alle matrici dei coefficienti. In questo caso, si fa riferimento alla regressione con termine di penalizzazione di tipo lasso. Il problema dei minimi quadrati penalizzati in questo caso è formulato come segue:

$$(\hat{A}_L, \hat{B}_L) = \underset{A, B}{\operatorname{arg\,min}} \left\{ \sum_{t=2}^T \|X_t - AX_{t-1}B'\|_F^2 + \lambda_{A_L} \|A\|_1 + \lambda_{B_L} \|B\|_1 \right\}, \quad (2.9)$$

dove $\|A\|_1 = \sum_{i,j}^m |a_{ij}|$, $\|B\|_1 = \sum_{i,j}^n |b_{ij}|$, $\lambda_{A_L} > 0$ e $\lambda_{B_L} > 0$.

\hat{A}_L e \hat{B}_L indicano rispettivamente le stime di A e B ottenuta mediante il metodo dei minimi quadrati con penalizzazione lasso. Per quanto riguarda λ_{A_L} e λ_{B_L} vale il discorso precedente relativo ai parametri di penalizzazione ridge. Si calcolano le derivate parziali e si pongono a zero:

$$\begin{aligned} \sum_{t=2}^T AX_{t-1}B'BX'_{t-1} - \sum_{t=2}^T X_tBX_{t-1} + \frac{\lambda_{A_L}}{2} \operatorname{sign}(A) &= 0, \\ \sum_{t=2}^T BX'_{t-1}A'AX'_{t-1} - \sum_{t=2}^T X'_tAX_{t-1} + \frac{\lambda_{B_L}}{2} \operatorname{sign}(B) &= 0. \end{aligned} \quad (2.10)$$

Sfortunatamente in questo caso non è possibile trovare una soluzione esplicita per A e B . Tuttavia è possibile ricorrere ad algoritmi numerici, come ad esempio l'algoritmo di discesa del gradiente, per potere calcolare \hat{A}_L e \hat{B}_L . Innanzitutto si utilizzano le matrici identità I_m e I_n come valore iniziale per \hat{A}_L e \hat{B}_L . Verrà applicata anche in questo caso una procedura di aggiornamento iterativa. Si aggiorna \hat{A}_L mediante l'algoritmo di discesa del gradiente ed utilizzando come valore di \hat{B}_L , la sua stima trovata al passo precedente. Dopodichè si utilizza \hat{A}_L appena aggiornato per aggiornare \hat{B}_L mediante l'algoritmo di discesa del gradiente e si continua fino a convergenza. Anche in questo caso si normalizza \hat{A}_L ad ogni passo.

Si osserva che gli stimatori presentati non impongono la condizione di causalità (1.0.1) durante la procedura di stima. Di conseguenza, le stime risultanti potrebbero non soddisfare necessariamente tale condizione, anche se il processo sottostante è stazionario e causale. Si spera che, se il processo è effettivamente causale, le consistenze degli stimatori assicurino che essi soddisfino la condizione di causalità con elevata probabilità.

La distribuzione asintotica dello stimatore $(\hat{A}_{LSE}, \hat{B}_{LSE})$ è la seguente. Siano

- $\alpha = \operatorname{vec}(A)$,
- $\beta = \operatorname{vec}(B')$,
- $\gamma = (\alpha', 0')' \in \mathbb{R}^{m^2+n^2}$,
- $W'_t = [(BX'_t) \otimes I, T \otimes (AX_t)]$,
- $H = \mathbb{E}(W_t W'_t) + \gamma \gamma'$,
- $\Xi_2 = H^{-1} \mathbb{E}(W_t \Sigma W'_t) H^{-1}$,

- $V = [\beta \otimes I, I \otimes \alpha]$

$$\sqrt{T} \begin{pmatrix} \text{vec}(\hat{A}_{LSE} - A) \\ \text{vec}(\hat{B}'_{LSE} - B') \end{pmatrix} \Rightarrow N(0, \Xi_2),$$

$$\sqrt{T} \left[\text{vec}(\hat{B}'_{LSE}) \otimes \text{vec}(\hat{A}_{LSE}) - \text{vec}(B') \otimes \text{vec}(A) \right] \Rightarrow N(0, V\Xi_2V').$$

Si utilizza la stessa distribuzione asintotica per (\hat{A}_R, \hat{B}_R) e (\hat{A}_L, \hat{B}_L) , utile in seguito al calcolo degli errori standard delle stime dei coefficienti.

Capitolo 3

Caso di studio

Il dataset in esame è costituito da osservazioni mensili dal 2001 al 2023 di quattro indici economici relativi a 10 paesi europei. In particolare si considerano: l'Indice di Volume della Produzione Industriale (VIP-I), relativo all'estrazione mineraria e di cava, alla produzione manifatturiera e alla fornitura di elettricità, gas, vapore e condizionamento dell'aria; l'Indice di Fatturato totale Industriale (FI), relativo relativo all'estrazione mineraria e di cava e alla produzione manifatturiera; l'Indice di Fatturato delle Vendite nel Commercio al Dettaglio esclusi veicoli a motore e motocicli (FD); l'Indice di Volume di Produzione nell'Edilizia (VIP-E). Tali indici sono relativi alle seguenti nazioni europee: Bulgaria (BG); Repubblica Ceca (CZ); Germania (DE); Ungheria (HU); Austria (AT); Polonia (PL); Romania (RO); Slovenia (SI); Slovacchia (SK); Svezia (SE). I dati sono stati recuperati dal sito di Eurostat [3], in particolare si sono selezionate le serie temporali già aggiustate stagionalmente. L'analisi è condotta sulle serie temporali mensili delle differenze prime logaritmiche. Si notano alcuni valori anomali nel grafico delle serie tem-

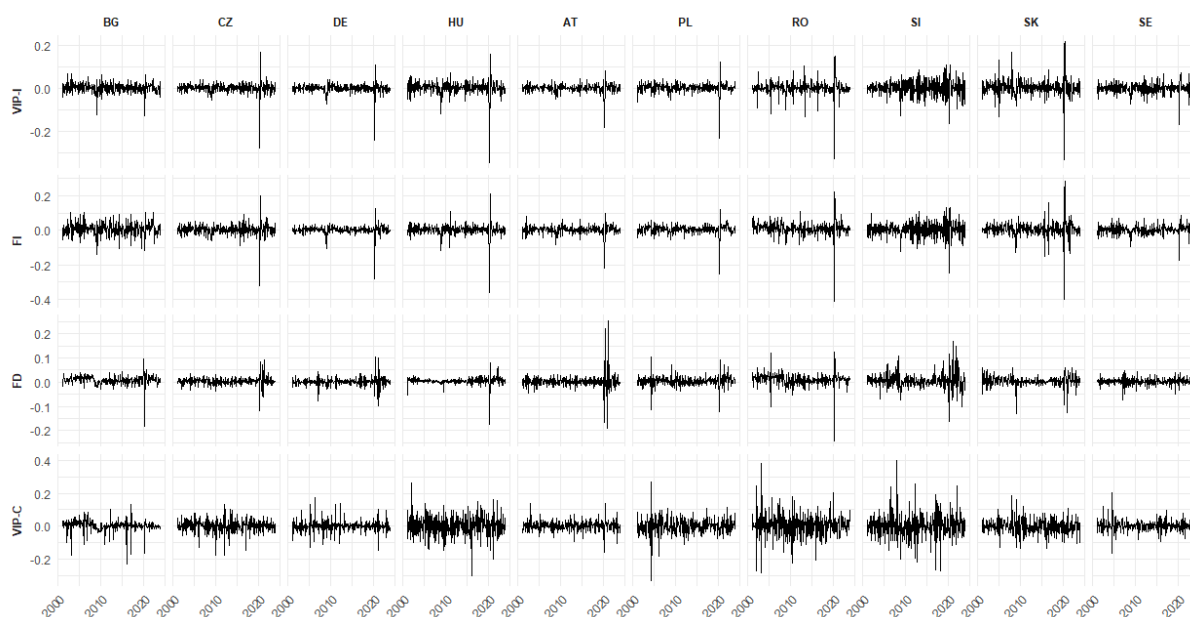


Figura 3.1: Serie temporali mensili delle differenze prime logaritmiche.

porali, in particolare verso il 2020. Ciò dovrebbe essere dovuto alla pandemia avvenuta a causa del Covid-19. Al fine di avere un confronto migliore tra le stime dei coefficienti per le varie serie, tutte le serie sono normalizzate in modo che la varianza di ciascun indicatore, quindi di ogni riga sia pari ad uno. In fase di adattamento del modello, si suddivide l'insieme dei dati in tre dataset: il set dei dati di stima composto dalle prime 108 osservazioni temporali è relativo alle osservazioni mensili da gennaio 2001 fino a dicembre 2009. Verrà utilizzato per trovare i valori ottimali dei parametri di penalizzazione delle stime ridge e lasso; il set dei dati di verifica composto da 156 osservazioni è relativo alle osservazioni mensili da gennaio 2010 fino a dicembre 2022. Verrà utilizzato per stimare le matrici dei coefficienti utilizzando

	VIP-I	FI	FD	VIP-C
VIP-I	0.607	-0.046	0.246	0.252
err.std.	0.101	0.191	0.111	0.122
t-value	5.979	-0.242	2.216	2.057
FI	0.446	0.055	0.252	0.262
err.std.	0.130	0.209	0.113	0.124
t-value	3.422	0.265	2.219	2.117
FD	0.176	0.239	0.178	0.098
err.std.	0.171	0.206	0.089	0.106
t-value	1.024	1.159	2.010	0.917
VIP-C	0.043	-0.002	0.016	0.202
err.std.	0.134	0.119	0.057	0.088
t-value	0.323	-0.016	0.283	2.295

Tabella 3.1: Stima coefficienti matrice \hat{A}_{LSE} .

i parametri di penalizzazione ottimali laddove servono; il set di dati di previsione composto da 12 osservazioni contiene le osservazioni mensili del 2023. Quest'ultimo verrà utilizzato in fase di previsione per valutare le capacità predittive dei vari modelli.

Innanzitutto, si stima il modello attraverso il metodo dei minimi quadrati senza penalizzazione. Osservando la Tabella 3.1, si nota che la prima colonna mostra l'effetto sugli attuali indici economici derivante dall'Indice di Volume della Produzione Industriale del mese precedente. L'impatto di crescita del VIP-I è statisticamente significativo al 95% per gli indici VIP-I e FI nel mese successivo, indicando che un aumento di tale indice implicherà un maggiore aumento degli indici attuali. La terza colonna mostra l'effetto sugli attuali indici economici derivante dalla crescita di FD del mese precedente. Gli effetti sono tutti positivi, tranne l'influenza sul VIP-C. La quarta colonna mostra l'effetto sugli attuali indici economici derivante dalla crescita di VIP-C del mese precedente. Gli effetti sono tutti positivi, tranne l'influenza su FD. Osservando le righe, invece, è possibile notare l'effetto di tutti gli indici del mese precedente sull'indice relativo alla riga considerata. La prima riga mostra che gli indici passati, ad eccezione di FI, hanno un'influenza positiva su VIP-I, e lo stesso si può per la seconda riga. Invece, considerando la terza e quarta riga, si nota che rispettivamente su FD e VIP-C ha effetto solamente l'indice stesso del mese precedente. Nella Tabella 3.2 sono riportate le stime dei coefficienti della matrice B . Le colonne rappresentano l'influenza degli indici del paese relativo nel mese precedente sugli indici di tutti i paesi nel mese attuale, mentre le righe mostrano l'effetto degli indici di tutti i paesi nel mese precedente sugli indici della nazione considerata. Si nota subito che i coefficienti statisticamente significativi al 95% sono pochi. La prima colonna mostra che l'influenza degli indici della Bulgaria nel mese precedente è positiva per Ungheria, Austria, Romania e Slovenia nel mese successivo. La terza colonna mostra che l'influenza degli indici della Germania nel mese precedente è negativa per Germania, Ungheria, Austria, Slovenia e Slovacchia nel mese successivo. Nella quinta colonna si osserva che l'influenza degli indici dell'Austria nel mese precedente è positiva per Ungheria, Romania, Slovenia e Slovacchia nel mese successivo. La nona colonna mostra che l'influenza degli indici della Slovacchia nel mese precedente è positiva e significativa solo la Slovenia nel mese successivo. Per quanto riguarda le colonne 4, 7, 8 e 10 si nota che gli indici economici del mese relativo a tali colonne sono influenzati unicamente dagli indici del paese stesso nel mese precedente. Per quanto riguarda le righe, gli indici dell'Ungheria sono influenzati positivamente da quelli di Bulgaria e Austria e negativamente da quelli di Germania e Ungheria stessa nel mese precedente. Gli indici della Romania sono influenzati positivamente da quelli di Bulgaria e Austria e negativamente da quelli della Bulgaria stessa nel mese precedente. Gli indici della Slovenia sono influenzati positivamente da quelli di Bulgaria, Austria e Slovacchia e negativamente da quelli di Germania e Slovenia stessa nel mese precedente. Gli indici della Slovacchia sono influenzati positivamente da quelli dell'Austria e negativamente da quelli della Germania nel mese precedente.

	BG	CZ	DE	HU	AT	PL	RO	SI	SK	SE
BG	-0.140	-0.033	-0.335	0.034	0.537	-0.283	-0.051	0.010	0.320	0.023
err.std.	0.253	0.313	0.237	0.345	0.300	0.270	0.378	0.423	0.412	0.181
t-value	-0.552	-0.106	-1.417	0.098	1.791	-1.049	-0.134	0.024	0.777	0.126
CZ	0.431	-0.483	-0.297	-0.192	0.612	-0.134	-0.184	0.032	0.428	-0.180
err.std.	0.304	0.386	0.278	0.419	0.337	0.316	0.455	0.517	0.514	0.232
t-value	1.419	-1.251	-1.069	-0.457	1.817	-0.424	-0.404	0.063	0.832	-0.777
DE	0.375	-0.015	-0.695	-0.136	0.706	-0.311	-0.113	-0.016	0.351	-0.113
err.std.	0.338	0.417	0.337	0.468	0.392	0.356	0.503	0.579	0.570	0.257
t-value	1.110	-0.037	-2.062	-0.292	1.800	-0.874	-0.225	-0.027	0.616	-0.439
HU	0.516	0.261	-0.437	-0.632	0.596	-0.111	-0.105	0.029	0.384	-0.206
err.std.	0.218	0.268	0.200	0.306	0.235	0.224	0.323	0.365	0.366	0.165
t-value	2.372	0.971	-2.187	-2.065	2.537	-0.495	-0.324	0.079	1.049	-1.248
AT	0.553	-0.205	-0.662	-0.026	0.216	-0.289	0.011	-0.090	0.403	-0.258
err.std.	0.283	0.350	0.288	0.412	0.287	0.304	0.454	0.455	0.511	0.219
t-value	1.952	-0.587	-2.298	-0.063	0.751	-0.951	0.024	-0.197	0.789	-1.181
PL	0.497	0.063	-0.357	-0.002	0.571	-0.390	-0.149	-0.011	0.208	-0.365
err.std.	0.299	0.367	0.277	0.402	0.327	0.319	0.446	0.498	0.499	0.218
t-value	1.663	0.173	-1.290	-0.005	1.745	-1.221	-0.333	-0.021	0.417	-1.674
RO	0.652	0.018	-0.232	-0.145	0.920	-0.201	-0.700	-0.027	0.422	-0.214
err.std.	0.166	0.206	0.151	0.228	0.182	0.171	0.259	0.279	0.282	0.125
t-value	3.925	0.089	-1.537	-0.637	5.050	-1.179	-2.705	-0.097	1.494	-1.711
SI	0.420	-0.330	-0.565	0.166	0.857	-0.210	-0.126	-0.676	0.478	-0.080
err.std.	0.141	0.177	0.128	0.194	0.155	0.146	0.210	0.258	0.236	0.104
t-value	2.976	-1.868	-4.401	0.858	5.541	-1.432	-0.600	-2.616	2.028	-0.765
SK	0.341	0.002	-0.464	-0.182	1.064	-0.140	-0.143	0.061	-0.002	-0.220
err.std.	0.215	0.265	0.202	0.294	0.240	0.219	0.321	0.361	0.347	0.158
t-value	1.591	0.008	-2.295	-0.619	4.442	-0.638	-0.445	0.169	-0.005	-1.391
SE	0.063	0.282	-0.367	0.073	0.343	0.054	-0.086	0.001	0.136	-0.532
err.std.	0.244	0.307	0.226	0.340	0.273	0.266	0.373	0.410	0.415	0.206
t-value	0.259	0.918	-1.625	0.214	1.258	0.202	-0.231	0.003	0.327	-2.579

Tabella 3.2: Stima coefficienti matrice \hat{B}_{LSE} .

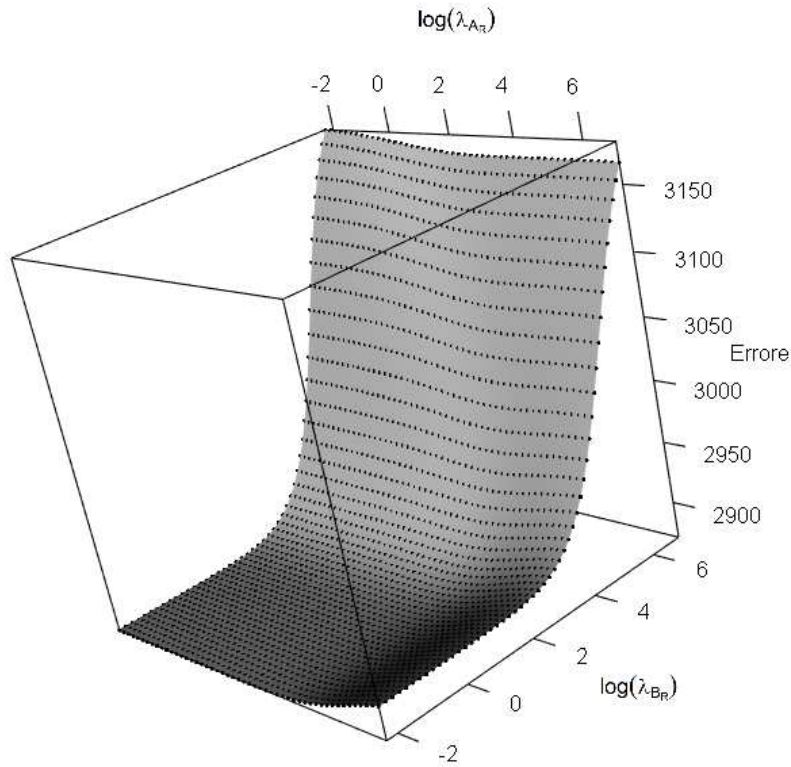


Figura 3.2: Errore di stima ridge al variare di $(\lambda_{A_R}, \lambda_{B_R})$.

Si procede ora con la fase di stima utilizzando il metodo dei minimi quadrati iterati con una penalizzazione di tipo ridge. Per prima cosa, viene definita una griglia di valori per i parametri di penalizzazione λ_{A_R} e λ_{B_R} . Si selezionano 50 valori per ciascun parametro, compresi tra 10^{-1} e 10^3 . Il modello viene adattato per ogni coppia di valori $(\lambda_{A_R}, \lambda_{B_R})$ e si sceglie la coppia ottimale. Non si sceglie la coppia che minimizza l'errore di stima, inteso come la somma degli errori di previsione al quadrato, ma si adotta un criterio più parsimonioso: si seleziona la coppia di valori associata all'errore di stima più elevato che risulta comunque inferiore all'errore minimo sommato alla deviazione standard degli errori di stima. La Figura (3.2) illustra la variazione dell'errore di stima al variare dei parametri di penalizzazione. Si nota che l'errore di stima aumenta all'aumentare di λ_{B_R} , mentre per λ_{A_R} non si osserva lo stesso andamento: al contrario, seppur di poco, nella parte alta del grafico, l'errore di stima diminuisce all'aumentare di λ_{A_R} . Questo potrebbe essere dovuto al fatto che, essendo la matrice A di dimensioni (4×4) , non sia necessario introdurre una penalizzazione per questa, mentre per B , contenente 100 elementi, sembra essere più sensato aggiungere una penalizzazione. Osservando il grafico (3.3) emerge che la coppia ottimale $(\lambda_{A_R}, \lambda_{B_R})$ (pallino rosso) è $(104.81, 86.85)$. La parte del grafico colorata in rosso indica il valore della somma tra l'errore di stima minimo e la deviazione standard di tutti gli errori di stima.

La Figura 3.4 mostra come varia l'errore di stima al variare di λ_{B_R} mantenendo λ_{A_R} costante pari al suo valore ottimale. Si osserva che l'errore di stima aumenta all'aumentare di λ_{B_R} . D'altra parte, dal grafico in Figura 3.5 si osserva che, una volta fissato il valore ottimale di λ_{B_R} , l'errore di stima rimane praticamente invariato al variare di λ_{A_R} . Questo conferma il fatto che, in questo contesto, non sia necessario penalizzare la matrice A .

Le Figure (3.6) e (3.7) mostrano rispettivamente come variano i coefficienti di \hat{B}_R e \hat{A}_R al variare di $(\lambda_{A_R}, \lambda_{B_R})$. Si osserva una contrazione dei coefficienti verso lo zero all'aumentare di λ_{B_R} , mentre questo non si verifica per λ_{A_R} . Il grafico (3.7) mostra anche che, nonostante la contrazione di alcuni coefficienti verso lo zero, la normalizzazione della matrice \hat{A}_R ad ogni passo dell'aggiornamento ricorsivo impedisce che ciò avvenga per tutti i coefficienti.

Infine, le Figure (3.8), (3.9), (3.10) e (3.11) mostrano come variano i coefficienti di \hat{A}_R e \hat{B}_R rispetto a un parametro di penalizzazione, mantenendo l'altro fisso. Queste figure confermano che l'applicazione di una penalizzazione agli elementi di B porta alla contrazione dei suoi coefficienti verso lo zero, mentre

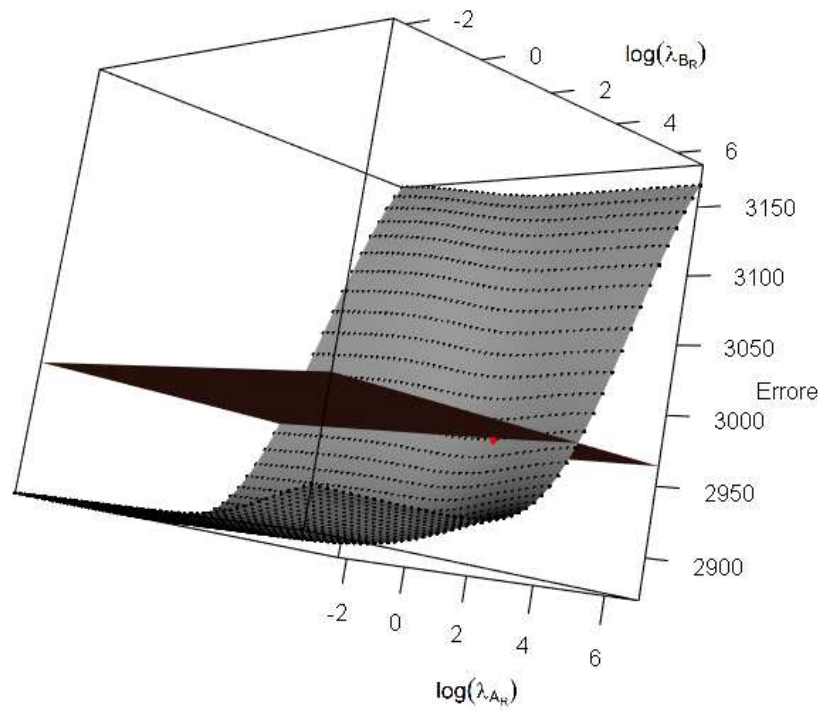


Figura 3.3: Errore di stima ridge al variare di $(\lambda_{A_R}, \lambda_{B_R})$ e valore ottimale $(\lambda_{A_R}, \lambda_{B_R})$.

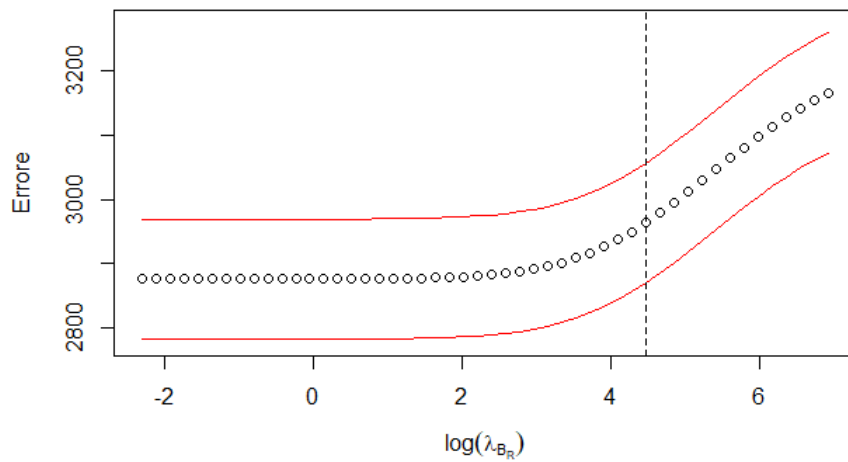


Figura 3.4: Errore di stima ridge al variare di λ_{B_R} con valore ottimale di λ_{A_R} .

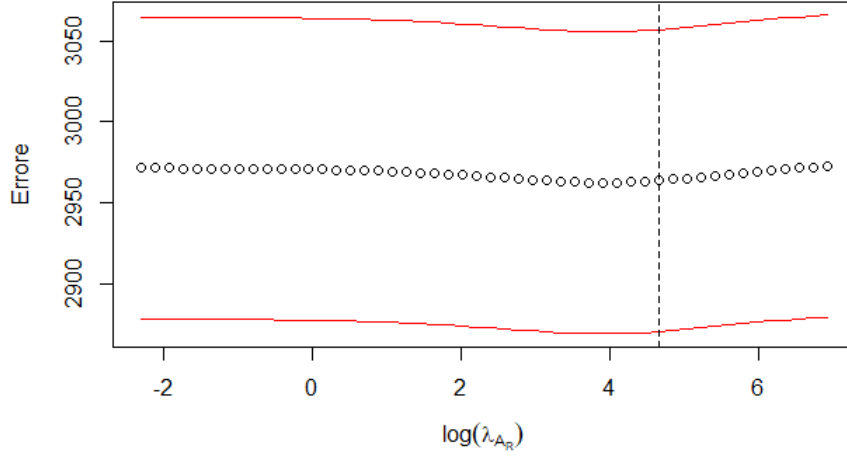


Figura 3.5: Errore di stima ridge al variare di λ_{A_R} con valore ottimale di λ_{B_R} .

	VIP-I	FI	FD	VIP-C
VIP-I	0.679	-0.164	0.300	0.067
err.std.	0.128	0.304	0.212	0.184
t-value	5.286	-0.538	1.413	0.366
FI	0.466	-0.033	0.318	0.077
err.std.	0.181	0.341	0.220	0.183
t-value	2.571	-0.098	1.445	0.419
FD	0.0948	0.116	0.186	0.027
err.std.	0.279	0.328	0.148	0.144
t-value	0.340	0.355	1.252	0.189
VIP-C	-0.008	0.001	0.036	0.183
err.std.	0.218	0.199	0.091	0.135
t-value	-0.038	0.006	0.391	1.348

Tabella 3.3: Stima coefficienti matrice \hat{A}_R .

l'applicazione di una penalizzazione ad A sembra non influire sulla stima dei coefficienti.

Si adotta quindi il metodo dei minimi quadrati penalizzati, applicando la penalizzazione esclusivamente sulla matrice B . Il valore ottimale per il parametro di penalizzazione λ_{B_R} emerge essere pari a 71.96, come rappresentato in Figura (3.12). Si presenta inoltre il grafico dei coefficienti stimati al variare di λ_{B_R} (Figura 3.13), per il quale si applicano le considerazioni effettuate precedentemente per il grafico (3.8). Nell'analisi dell'andamento della matrice \hat{A}_R , si osserva che solamente due coefficienti risultano statisticamente significativi. L'impatto di crescita del VIP-I è statisticamente significativo al 95% per gli indici VIP-I e FI nel mese successivo, come accade nel caso della stima senza termine di penalizzazione. Per quanto riguarda \hat{B}_R , si ha un solo coefficiente statisticamente significativo tale per cui gli indici dell'Austria influenzano positivamente quelli della Romania nel mese successivo.

In seguito, si procede con la stima dei parametri A e B mediante una penalizzazione di tipo lasso, applicata esclusivamente alla matrice B , per le stesse considerazioni espresse nella fase di stima ridge. Per la procedura iterativa di stima, si è adottato l'algoritmo di discesa del gradiente. Tuttavia, al fine di mitigare il costo computazionale in termini di tempo, il numero di iterazioni è stato limitato. Le curve frastagliate presenti in Figura 3.21 e Figura 3.15 sono conseguenza del numero esigui di iterazioni applicate nell'implementazione del codice per la stima dei parametri. In particolare, ad ogni iterazione si alternano una procedura di discesa del gradiente moderata e un processo di aggiornamento iterativo. Il valore ottimale per il parametro λ_{B_L} risulta essere 13.26.

Le stime dei coefficienti delle matrici \hat{A}_L e \hat{B}_L sono presentate in Tabella 3.5 e Tabella 3.6. Si osserva che l'impatto del VIP-I è statisticamente significativo al 95% per gli indici VIP-I, FI e FD nel mese

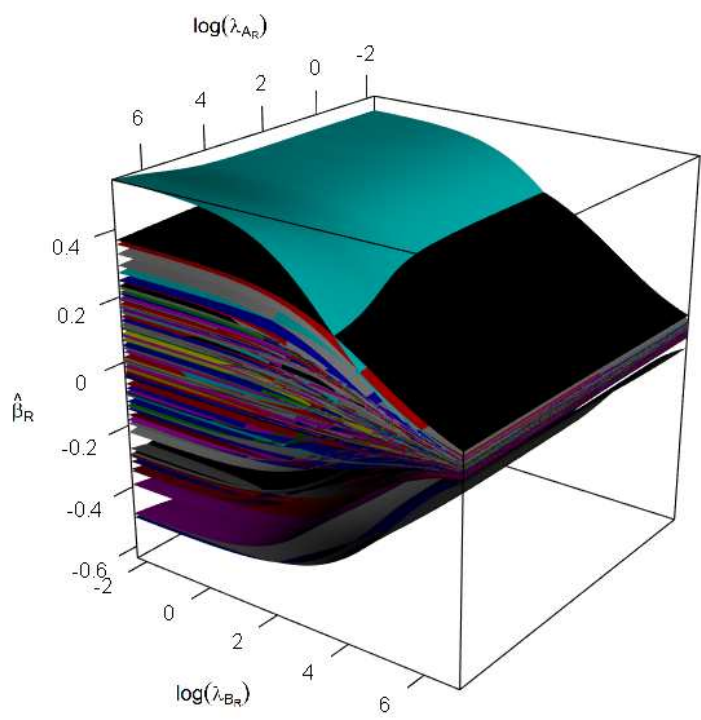


Figura 3.6: Coefficienti matrice \hat{B}_R al variare di $(\lambda_{A_R}, \lambda_{B_R})$.

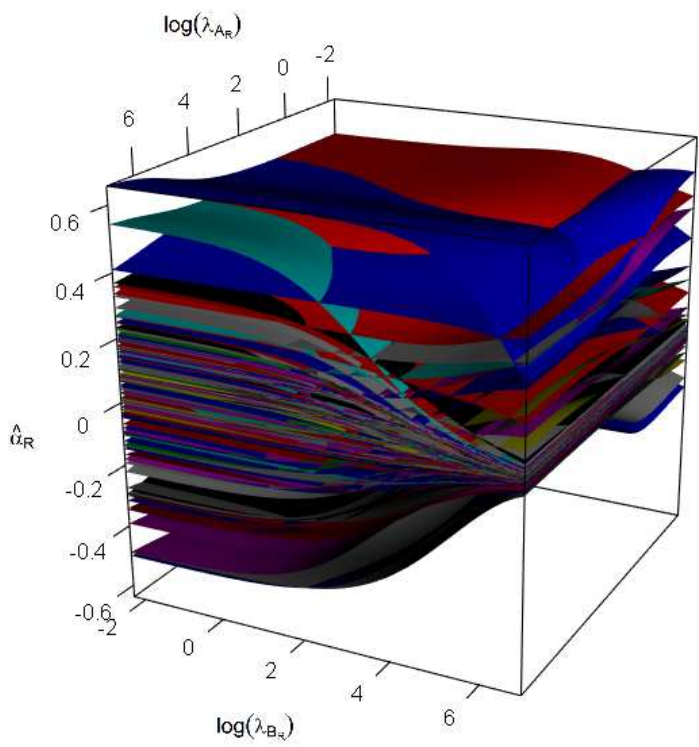


Figura 3.7: Coefficienti matrice \hat{A}_R al variare di $(\lambda_{A_R}, \lambda_{B_R})$.

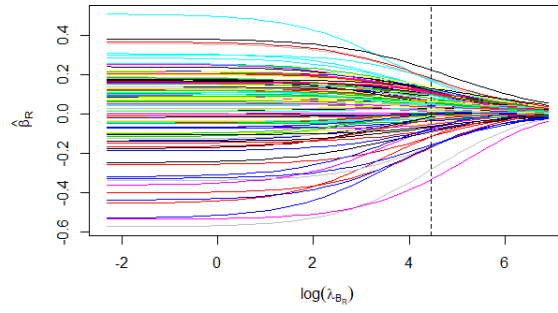


Figura 3.8: Coefficienti matrice \hat{B}_R al variare di λ_{B_R} con valore ottimale di λ_{A_R} .

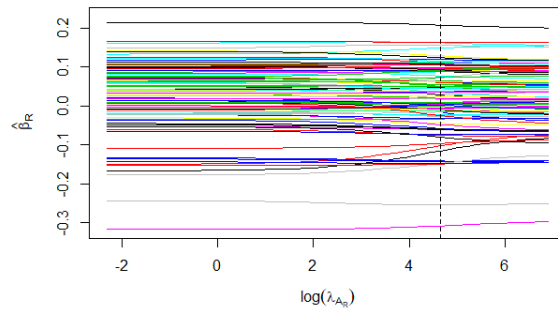


Figura 3.9: Coefficienti matrice \hat{B}_R al variare di λ_{A_R} con valore ottimale di λ_{B_R} .

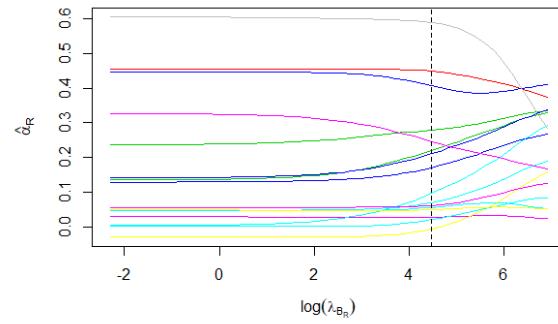


Figura 3.10: Coefficienti matrice \hat{A}_R al variare di λ_{B_R} con valore ottimale di λ_{A_R} .

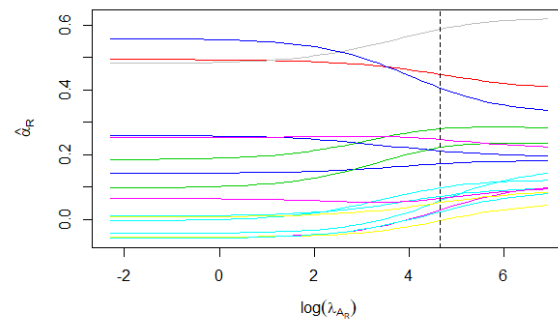


Figura 3.11: Coefficienti matrice \hat{A}_R al variare di λ_{A_R} con valore ottimale di λ_{B_R} .

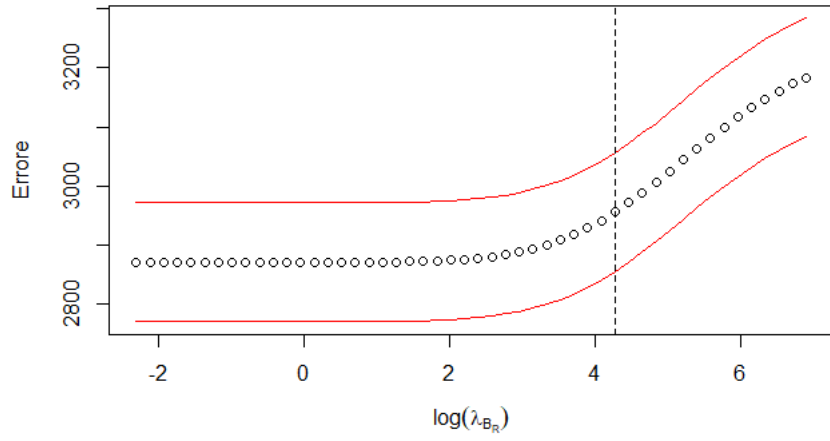


Figura 3.12: Errore di stima ridge al variare di λ_{B_R} .

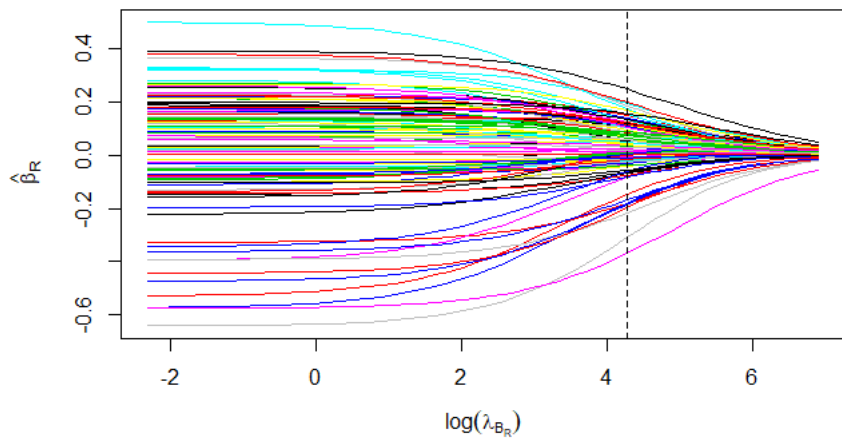


Figura 3.13: Coefficienti matrice \hat{B}_R al variare di λ_{B_R} .

	BG	CZ	DE	HU	AT	PL	RO	SI	SK	SE
BG	-0.034	0.053	-0.013	-0.068	0.189	-0.035	-0.070	0.070	0.179	-0.021
err.std.	0.282	0.356	0.269	0.393	0.308	0.297	0.421	0.489	0.478	0.228
t-value	-0.121	0.149	-0.049	-0.173	0.615	-0.118	-0.167	0.144	0.373	-0.093
CZ	0.189	-0.049	-0.014	-0.245	0.265	-0.060	-0.169	0.133	0.241	-0.104
err.std.	0.353	0.453	0.334	0.501	0.371	0.371	0.529	0.620	0.611	0.294
t-value	0.536	-0.108	-0.042	-0.489	0.715	-0.161	-0.319	0.215	0.394	-0.354
DE	0.170	0.113	-0.145	-0.188	0.305	-0.126	-0.131	0.088	0.225	-0.113
err.std.	0.384	0.474	0.361	0.525	0.414	0.396	0.569	0.659	0.641	0.303
t-value	0.442	0.237	-0.401	-0.358	0.736	-0.318	-0.230	0.133	0.350	-0.372
HU	0.249	0.304	-0.020	-0.524	0.315	-0.044	-0.106	0.122	0.255	-0.154
err.std.	0.279	0.361	0.267	0.423	0.292	0.286	0.424	0.486	0.478	0.229
t-value	0.893	0.843	-0.075	-1.239	1.077	-0.152	-0.249	0.251	0.533	-0.671
AT	0.216	0.039	-0.133	-0.131	-0.002	-0.094	-0.033	0.021	0.174	-0.117
err.std.	0.294	0.383	0.322	0.473	0.286	0.327	0.505	0.470	0.562	0.247
t-value	0.736	0.102	-0.412	-0.277	-0.005	-0.287	-0.066	0.045	0.310	-0.472
PL	0.233	0.188	-0.046	-0.082	0.238	-0.162	-0.144	0.089	0.129	-0.229
err.std.	0.341	0.426	0.319	0.468	0.360	0.361	0.505	0.587	0.573	0.273
t-value	0.685	0.442	-0.145	-0.174	0.660	-0.448	-0.286	0.151	0.226	-0.837
RO	0.310	0.266	0.069	-0.207	0.425	-0.058	-0.541	0.123	0.307	-0.135
err.std.	0.193	0.244	0.183	0.271	0.201	0.199	0.312	0.333	0.333	0.158
t-value	1.610	1.087	0.378	-0.762	2.120	-0.291	-1.735	0.368	0.921	-0.852
SI	0.121	0.074	-0.068	-0.074	0.263	-0.072	-0.024	-0.446	0.315	0.023
err.std.	0.167	0.224	0.166	0.248	0.170	0.179	0.259	0.319	0.293	0.137
t-value	0.727	0.332	-0.409	-0.297	1.544	-0.401	-0.091	-1.396	1.074	0.167
SK	0.170	0.208	-0.033	-0.263	0.499	0.000	-0.148	0.163	0.003	-0.186
err.std.	0.259	0.331	0.255	0.368	0.271	0.266	0.398	0.459	0.431	0.217
t-value	0.657	0.629	-0.128	-0.715	1.843	0.000	-0.372	0.356	0.006	-0.858
SE	0.020	0.182	-0.058	0.012	0.144	0.073	-0.066	0.067	0.121	-0.316
err.std.	0.281	0.364	0.274	0.404	0.299	0.312	0.429	0.492	0.493	0.268
t-value	0.071	0.499	-0.210	0.030	0.482	0.234	-0.154	0.137	0.245	-1.180

Tabella 3.4: Stima coefficienti matrice \hat{B}_R .

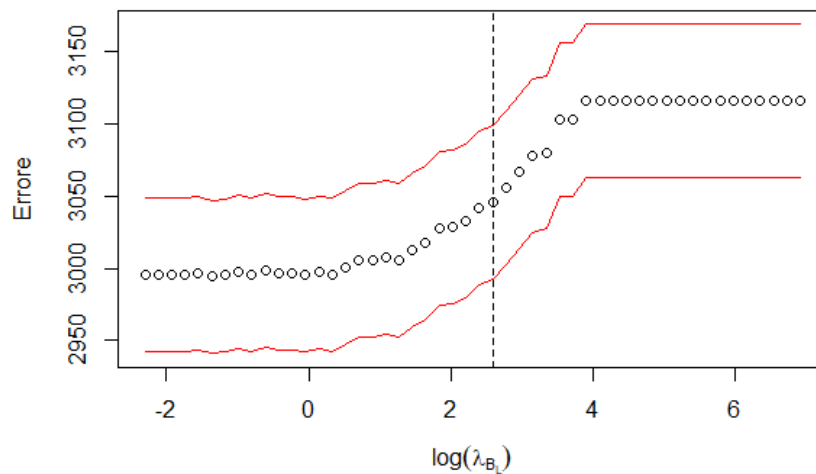


Figura 3.14: Errore di stima lasso al variare di λ_{BL} .

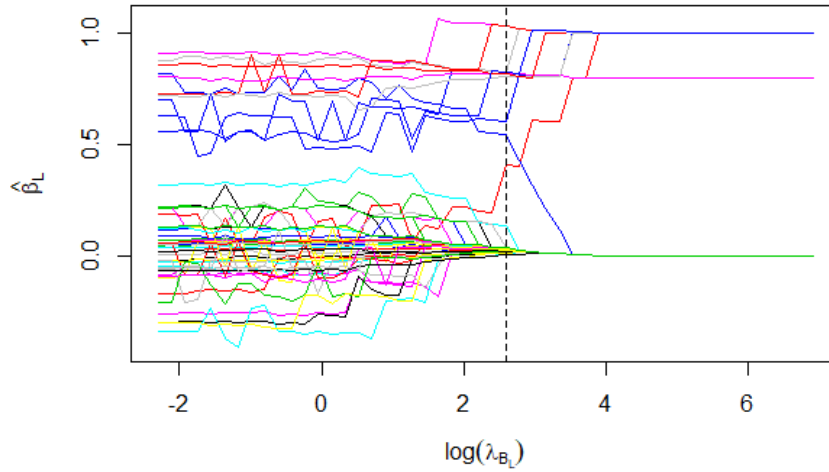


Figura 3.15: Coefficienti matrice \hat{B}_L al variare di λ_{B_L} .

	VIP-I	FI	FD	VIP-C
VIP-I	-0.334	0.038	0.483	0.034
err.std.	0.119	0.150	0.098	0.111
t-value	-2.812	0.252	4.922	0.306
FI	-0.308	-0.004	0.499	0.009
err.std.	0.123	0.155	0.100	0.113
t-value	-2.499	-0.028	4.973	0.078
FD	-0.409	0.187	0.233	0.092
err.std.	0.101	0.123	0.076	0.102
t-value	-4.031	1.525	3.072	0.899
VIP-C	-0.005	0.025	0.080	-0.190
err.std.	0.099	0.088	0.059	0.081
t-value	-0.052	0.285	1.352	-2.349

Tabella 3.5: Stima coefficienti matrice \hat{A}_L .

successivo. Tuttavia a differenza delle due stime precedenti, in questo caso l'influenza è negativa. La terza colonna, come nel caso della stima LSE senza penalizzazione, mostra effetto positivo sugli attuali indici economici derivante dalla crescita di FD del mese precedente, tranne che sul VIP-C. Per quanto concerne le stime della matrice B , si evince che gli indici della Bulgaria nel mese precedente influenzano positivamente gli indici degli altri paesi ad eccezione della Germania, Polonia e Svezia. Gli indici della Svezia influenzano positivamente quelli dell'Austria, Romania, Slovenia e della Svezia stessa nel mese successivo.

Al fine di confrontare le prestazioni del modello MAR(1) rispetto a quelle del modello VAR(1) tradizionale, è stato stimato anche un modello VAR(1) mediante la vettorizzazione della matrice dei dati, espresso come:

$$vec(X_t) = \Phi vec(X_{t-1}) + E_t.$$

I valori dell'Errore Quadratico Medio (MSE) per i quattro modelli sono presentati nella Tabella 3.7. Emerge un notevole divario tra l'MSE del modello VAR(1) e quello del modello MAR(1), dovuto al maggior numero di parametri stimati nel modello VAR(1), 1600 rispetto ai 116 del modello MAR(1). Tuttavia, analizzando la Tabella 3.8, che illustra l'MSE di previsione calcolato utilizzando le previsioni derivanti dal set di dati per le previsioni, si nota che l'MSE del modello VAR(1) risulta essere il più alto, indicando una maggiore precisione predittiva del modello MAR(1).

In conclusione è stato calcolato l'indice R-quadro medio per i tre modelli stimati. Come indicato nella

	BG	CZ	DE	HU	AT	PL	RO	SI	SK	SE
BG	0.728	0.077	0.015	-0.010	0.029	0.075	0.286	0.230	0.021	0.261
err.std.	0.295	0.340	0.268	0.395	0.312	0.298	0.421	0.436	0.439	0.233
t-value	2.467	0.226	0.055	-0.026	0.092	0.252	0.681	0.528	0.049	1.120
CZ	0.788	0.718	0.009	-0.060	0.041	0.078	0.010	0.203	-0.083	0.277
err.std.	0.350	0.428	0.317	0.488	0.375	0.364	0.531	0.552	0.549	0.283
t-value	2.249	1.679	0.029	-0.124	0.110	0.215	0.018	0.367	-0.152	0.978
DE	0.781	0.198	0.145	-0.061	0.259	0.054	0.009	0.110	-0.103	0.075
err.std.	0.279	0.338	0.258	0.395	0.301	0.293	0.431	0.447	0.453	0.231
t-value	2.794	0.584	0.560	-0.154	0.861	0.183	0.020	0.246	-0.228	0.326
HU	0.792	0.522	-0.194	0.455	0.273	0.004	-0.049	0.135	-0.156	0.403
err.std.	0.305	0.372	0.283	0.452	0.341	0.326	0.482	0.508	0.482	0.253
t-value	2.601	1.401	-0.685	1.008	0.802	0.013	-0.101	0.266	-0.323	1.597
AT	0.699	0.351	-0.048	0.066	0.002	0.038	0.022	0.209	-0.175	0.396
err.std.	0.201	0.246	0.199	0.291	0.211	0.211	0.314	0.318	0.337	0.168
t-value	3.474	1.426	-0.239	0.226	0.011	0.180	0.072	0.658	-0.519	2.360
PL	0.760	0.285	-0.063	-0.147	0.022	0.400	0.082	0.213	-0.071	0.372
err.std.	0.373	0.453	0.345	0.531	0.398	0.398	0.575	0.609	0.605	0.313
t-value	2.036	0.629	-0.183	-0.276	0.054	1.007	0.143	0.350	-0.118	1.189
RO	0.872	0.558	-0.042	-0.392	0.190	0.008	0.284	0.259	-0.200	0.590
err.std.	0.257	0.305	0.234	0.356	0.291	0.270	0.391	0.409	0.393	0.209
t-value	3.388	1.829	-0.177	-1.103	0.654	0.031	0.726	0.632	-0.509	2.819
SI	0.684	0.533	0.034	-0.099	-0.287	0.119	0.129	0.748	-0.383	0.661
err.std.	0.172	0.202	0.155	0.235	0.182	0.176	0.257	0.285	0.270	0.139
t-value	3.976	2.640	0.222	-0.422	-1.583	0.674	0.502	2.629	-1.418	4.758
SK	0.792	0.479	0.084	-0.347	0.280	0.022	-0.009	0.336	0.000	0.335
err.std.	0.265	0.316	0.244	0.369	0.296	0.277	0.408	0.425	0.420	0.215
t-value	2.995	1.515	0.346	-0.941	0.947	0.081	-0.022	0.791	0.000	1.556
SE	0.574	0.017	-0.079	-0.062	0.128	0.006	0.257	0.049	0.006	0.685
err.std.	0.309	0.366	0.283	0.428	0.345	0.323	0.474	0.491	0.485	0.259
t-value	1.857	0.046	-0.280	-0.145	0.370	0.018	0.543	0.101	0.012	2.643

Tabella 3.6: Stima coefficienti matrice \hat{B}_L .

Modello	MAR(1)	MAR(1) Ridge	MAR(1) Lasso	VAR(1)
MSE	6169.51	6402.22	6207.43	3620.10

Tabella 3.7: MSE per ciascun modello adattato all'insieme di verifica.

Modello	MAR(1)	MAR(1) Ridge	MAR(1) Lasso	VAR(1)
MSE	272.10	256.89	258.91	287.10

Tabella 3.8: MSE di previsione per ciascun modello.

Tabella 3.9, i risultati evidenziano valori significativamente bassi per l'indice R-quadro medio. Questa osservazione suggerisce che i modelli analizzati non si adattano in maniera efficace ai dati.

Stima	LSE	LSE Ridge	LSE Lasso
R^2 medio	0.176	0.145	0.171

Tabella 3.9: Indice R^2 medio modello MAR(1) per ciascun metodo di stima.

Si riportano anche i grafici ACF e temporali delle serie dei residui. Si può notare come i residui sembrano presentare lieve autocorrelazione e come il grafico temporale dei residui non indichi un andamento casuale attorno allo zero.

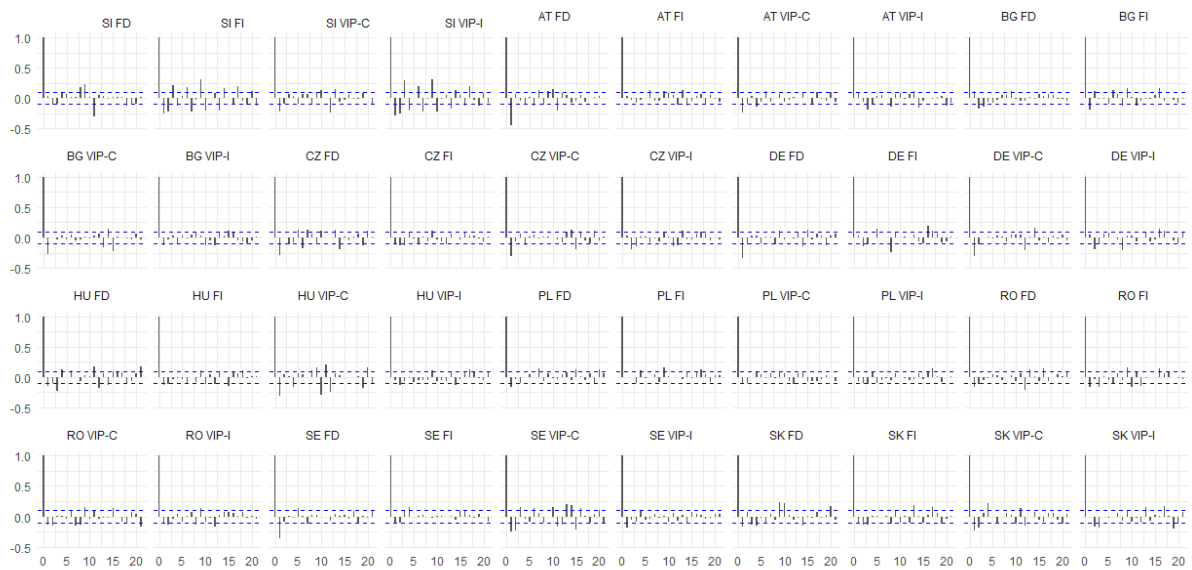


Figura 3.16: Funzione di Autocorrelazione degli errori con stimatori \hat{A}_{LSE} e \hat{B}_{LSE} .

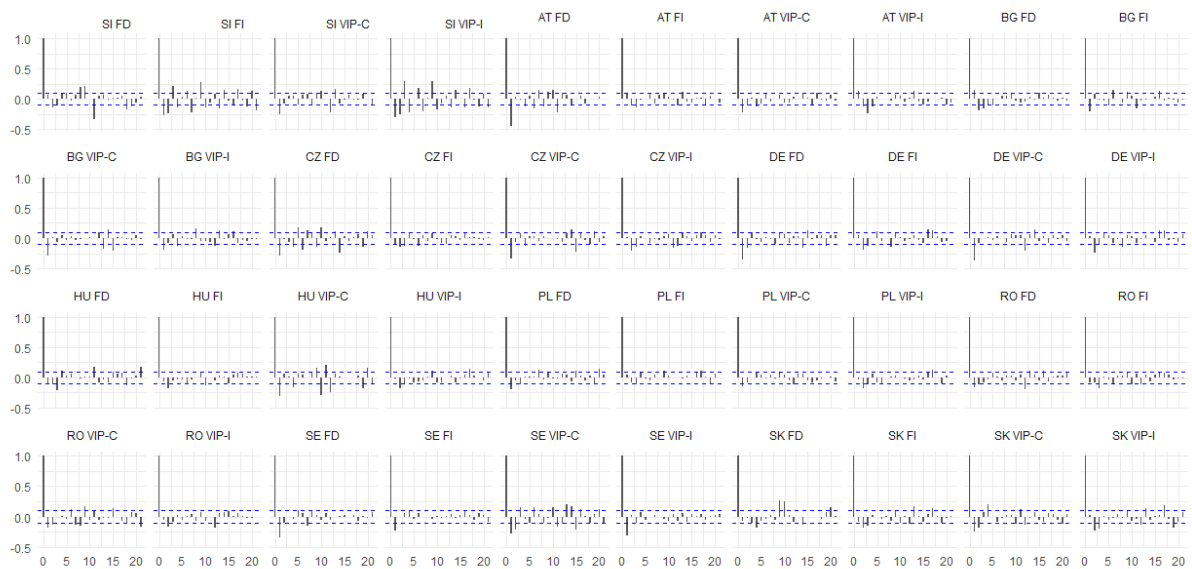


Figura 3.17: Funzione di Autocorrelazione degli errori con stimatori \hat{A}_R e \hat{B}_R .

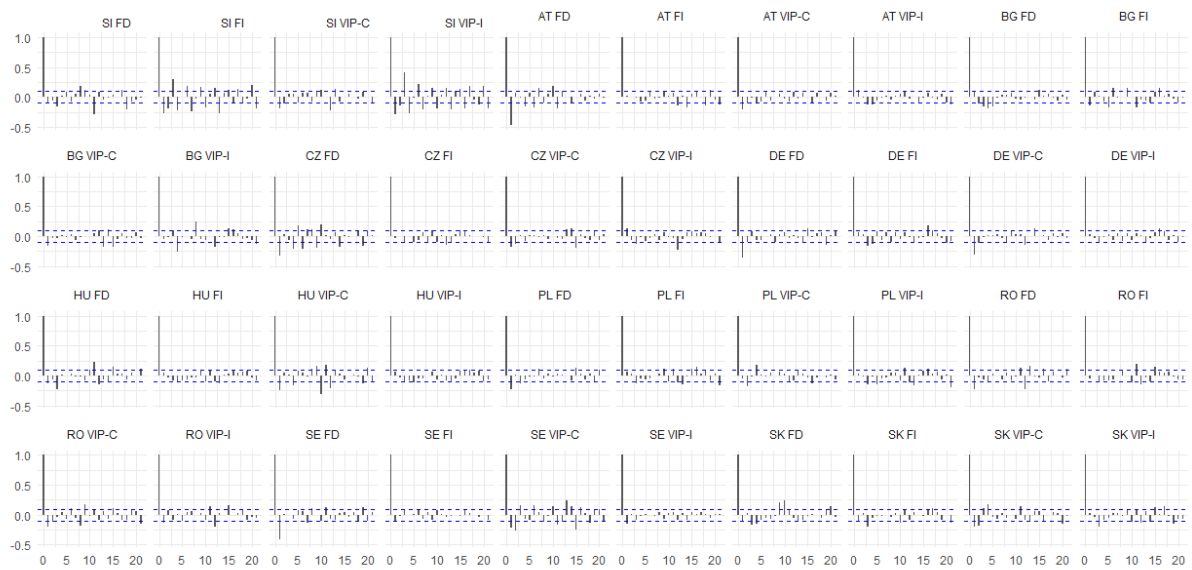


Figura 3.18: Funzione di Autocorrelazione degli errori con stimatori \hat{A}_L e \hat{B}_L .

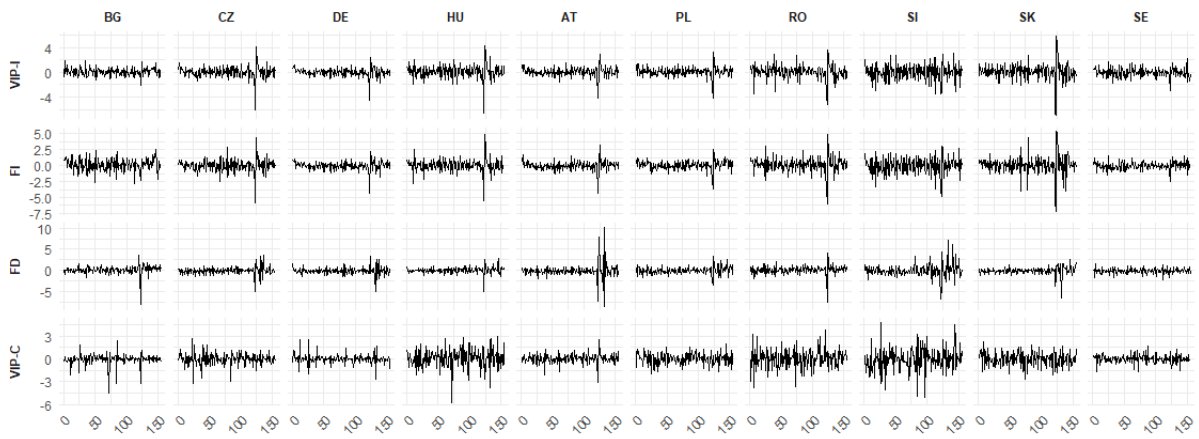


Figura 3.19: Grafici degli errori con stimatori \hat{A}_{LSE} e \hat{B}_{LSE} .

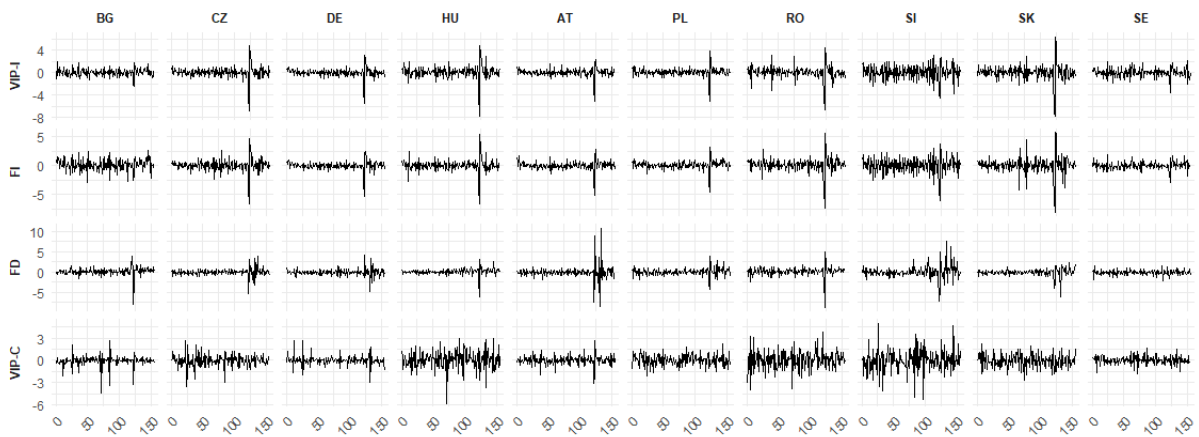


Figura 3.20: Grafici degli errori degli errori con stimatori \hat{A}_R e \hat{B}_R .

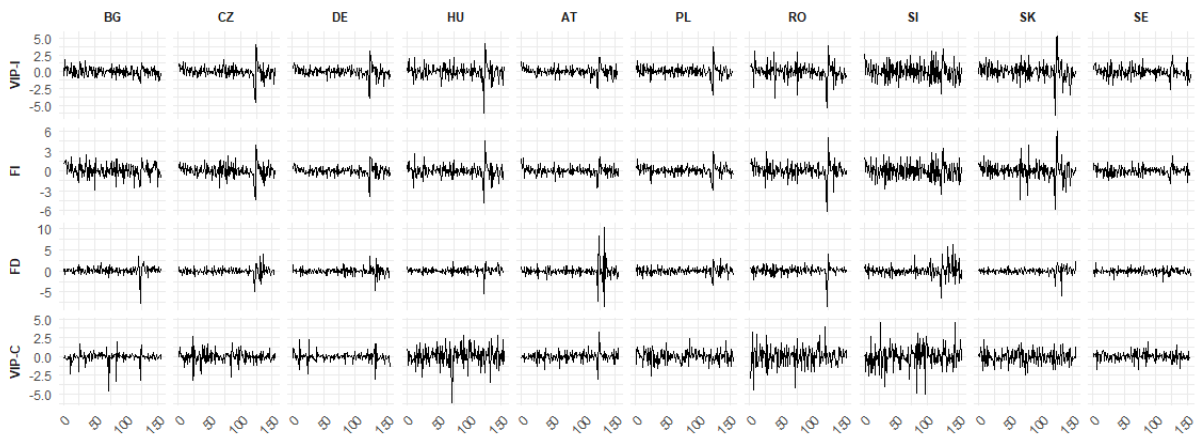


Figura 3.21: Grafici degli errori degli errori con stimatori \hat{A}_L e \hat{B}_L .

Conclusioni

Dall'analisi effettuata emerge che, sebbene il modello $MAR(1)$ mostri una capacità predittiva superiore rispetto al modello $VAR(1)$ e necessiti di un numero di gran lunga inferiore di parametri da stimare, si riscontrano delle difficoltà nell'adattamento ai dati osservati. Questa limitazione potrebbe essere attribuita a diversi fattori, tra cui l'alta dimensionalità dei dati, che può comportare complessità aggiuntiva e aumentare il rischio di sovradattamento. L'introduzione di stimatori basati sui minimi quadrati con l'aggiunta di termini di penalizzazione si rivela promettente per migliorare le performance del modello. Questi approcci consentono di bilanciare la precisione della previsione con la complessità del modello, riducendo così il rischio di sovradattamento e migliorando la generalizzazione ai dati non osservati. I risultati dell'errore quadratico medio di previsione (MSE) riportati nella Tabella 3.8 confermano l'efficacia di questo approccio nel migliorare le prestazioni predittive. Inoltre, una possibile strategia per migliorare ulteriormente il modello potrebbe essere l'aumento dell'ordine autoregressivo. Ampliando il numero di ritardi considerati, il modello potrebbe catturare meglio le dinamiche complesse presenti nei dati e migliorare la capacità di previsione su periodi temporali più lunghi. Tuttavia, è importante bilanciare questa espansione con il rischio di sovrastimare la complessità del modello e quindi compromettere le prestazioni predittive su dati futuri. In conclusione, l'utilizzo di stimatori penalizzati e l'esplorazione di modelli autoregressivi di ordine superiore potrebbero rappresentare due approcci chiave per affrontare le sfide nell'adattamento dei modelli $MAR(1)$ ai dati reali. Queste strategie offrono opportunità significative per migliorare la capacità predittiva e la capacità di generalizzazione dei modelli autoregressivi per serie temporali a valori matriciali.

Bibliografia

- [1] Alessandro Celani and Paolo Pagnottoni. Matrix autoregressive models: Generalization and bayesian estimation. *Studies in Nonlinear Dynamics & Econometrics*, (0), 2023.
- [2] Rong Chen, Han Xiao, and Dan Yang. Autoregressive models for matrix-valued time series. *Journal of Econometrics*, 222(1):539–560, 2021.
- [3] Eurostat. Short-term business statistics (sts). Recuperato da <https://ec.europa.eu/eurostat/data>. 2024.
- [4] Nan-Jung Hsu, Hsin-Cheng Huang, and Ruey S Tsay. Matrix autoregressive spatio-temporal models. *Journal of Computational and Graphical Statistics*, 30(4):1143–1155, 2021.
- [5] Gary C McDonald. Ridge regression. *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(1):93–100, 2009.
- [6] Jonas Ranstam and Jonathan A Cook. Lasso regression. *Journal of British Surgery*, 105(10):1348–1348, 2018.
- [7] Hong-Fan Zhang. Additive autoregressive models for matrix valued time series. *Journal of Time Series Analysis*, 2023.