

出國報告（出國類別：實習）

參加 SEACEN 研訓中心舉辦之
「貨幣與金融穩定政策預測」課程
研習報告書

服務機關：中央銀行

姓名職稱：朱浩榜/經濟研究處三等專員

派赴國家：尼泊爾

出國期間：112 年 3 月 26 日至 4 月 2 日

報告日期：112 年 6 月

目 錄

壹、前言.....	1
貳、大量變數的預測模型	2
一、正規化迴歸.....	2
二、貝氏 VAR.....	6
三、動態因子模型.....	10
參、經濟預測的相關議題	12
一、預測組合.....	12
二、結構改變.....	14
肆、應用範例：預測台灣的經濟成長率	17
一、簡介.....	17
二、資料說明.....	18
三、預測方法.....	20
四、實證結果.....	22
伍、結論與建議	25
參考資料.....	26

參加 SEACEN 研訓中心舉辦之 「貨幣與金融穩定政策預測」課程研習報告書

壹、前言

東南亞國家中央銀行聯合會（South East Asian Central Banks, SEACEN）所屬研訓中心於 2023 年 3 月在尼泊爾加德滿都舉辦「貨幣與金融穩定政策預測」（Forecasting for Monetary and Financial Stability Policies）課程，課程內容涵蓋大量變數之預測模型，以及有關經濟預測之常見議題。

本次研習參加學員共 26 人，分別來自我國與菲律賓、印尼、馬來西亞、泰國、緬甸、印度、尼泊爾等 8 國之中央銀行，以及尼泊爾之財政部、國家計畫委員會（National Planning Commission）與國家統計局（National Statistics Office）等主辦國之政府機關。課程講師係由 SEACEN 研訓中心、瑞典厄勒布魯大學（Örebro University）之經濟學家與研究人員擔任，學識及經驗均相當豐富；主辦單位另亦邀請美國聯邦準備理事會（Federal Reserve Board of Governors）之經濟學家，透過網路視訊分享其研究成果。

本報告共分為 5 個部分。除前言外，第 2 部分簡介常見的大量變數預測模型，包括正規化迴歸（regularized regression）、貝氏（Bayesian）向量自我迴歸（vector autoregression, VAR）與動態因子模型（dynamic factor model）等。第 3 部分討論預測組合（forecast combination）、結構改變（structural break）等經濟預測之相關議題。第 4 部分則應用前述介紹之內容對台灣經濟成長率進行預測，以作為說明範例。最後，第 5 部分為結論與建議¹。

¹ 本報告內容主要參考 SEACEN 研習講義（2023）。

貳、大量變數的預測模型

中央銀行藉由貨幣或金融穩定政策達成其政策目標，惟無論貨幣或金融穩定政策，自施行至發揮作用往往需要一段時間，故在未來具有不確定性的情況下，經濟預測便扮演重要角色。具體而言，經濟預測係指「提供某變數未來數值之相關資訊的過程」，而如何應用現有資訊達成此目的，通常須借助預測模型。直覺上，愈多變數或可提供愈多資訊；但若欲同時納入大量變數，則維度增加反將降低自由度並影響模型的參數估計。有鑑於此，文獻上提出許多應用大量變數的預測模型，以解決維度過高的問題。以下依序介紹正規化迴歸、貝氏 VAR 及動態因子模型等較為常見的大量變數預測模型。

一、正規化迴歸

假設線性迴歸模型如下：

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2).$$

其中， \mathbf{y} 為 $n \times 1$ 的被解釋變數向量， \mathbf{X} 為 $n \times k$ 的解釋變數矩陣， $\boldsymbol{\epsilon}$ 為常態分配之殘差項，而 $\boldsymbol{\beta}$ 為待估參數。常用的一般最小平方法（ordinary least squares, OLS）係極小化損失函數（loss function）

$$\mathcal{L}_{\text{ols}} = \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2, \quad (1.1)$$

並得到估計係數及誤差、變異數分別為

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}, \\ \text{Bias}(\hat{\boldsymbol{\beta}}) &= \mathbb{E}(\hat{\boldsymbol{\beta}}) - \boldsymbol{\beta}, \\ \text{Var}(\hat{\boldsymbol{\beta}}) &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}. \end{aligned}$$

其中，

$$\sigma^2 = \frac{\epsilon' \epsilon}{n - k} .$$

由上可知，當解釋變數眾多時 ($k \rightarrow n$)， σ^2 及 $\text{Var}(\hat{\boldsymbol{\beta}})$ 將隨之上升，此或肇因於多重共線性 (multicollinearity) 之故；而當解釋變數過多時 ($k > n$)， $(\mathbf{X}'\mathbf{X})$ 將不可逆 (not invertible)，使得 $\hat{\boldsymbol{\beta}}$ 不具唯一解。

因此，在解釋變數眾多的情況下，可在線性迴歸模型中加入某些懲罰項 (penalty term)。亦即，對迴歸係數加上某些限制 (constraints) 使其向 0 縮減 (shrink to zero)，以避免參數過多的問題，此即「正規化迴歸」的概念。正規化迴歸屬於監督式 (supervised) 的機器學習 (machine learning) 方法，依據懲罰項的設定方式，可分為脊迴歸 (ridge regression)、LASSO (least absolute shrinkage and selection operator) 與 Elastic Net (或稱「彈性網路」) 等幾種做法。

(一) 脊迴歸

若在 (1.1) 式之 OLS 損失函數加上懲罰項 $\lambda \|\boldsymbol{\beta}\|^2$ 以控制係數 (Hoerl and Kennard, 1970)，則為脊迴歸之損失函數：

$$\mathcal{L}_{\text{ridge}} = \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda \|\boldsymbol{\beta}\|^2. \quad (1.2)$$

其中， $\lambda \|\boldsymbol{\beta}\|^2$ 亦稱為 L2 規範懲罰項 (L2 norm penalty term)， λ 為調控參數 (tuning parameter)，用於控制係數。透過 (1.2) 式之損失函數，可得到脊迴歸的估計係數及誤差、變異數分別為

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{\text{ridge}} &= (\mathbf{X}'\mathbf{X} + \mathbf{I}\lambda)^{-1} \mathbf{X}'\mathbf{y}, \\ \text{Bias}(\hat{\boldsymbol{\beta}}_{\text{ridge}}) &= (\mathbf{X}'\mathbf{X} + \mathbf{I}\lambda)^{-1} \boldsymbol{\beta}, \\ \text{Var}(\hat{\boldsymbol{\beta}}_{\text{ridge}}) &= \sigma^2 (\mathbf{X}'\mathbf{X} + \mathbf{I}\lambda)^{-1} \mathbf{X}'\mathbf{X} (\mathbf{X}'\mathbf{X} + \mathbf{I}\lambda)^{-1}. \end{aligned}$$

由此可知，當 $\lambda = 0$ 時懲罰項為 0，此時脊迴歸的估計結果與 OLS 相

同；當 $\lambda \rightarrow \infty$ 時懲罰項甚大，此時模型係數將趨近於 0。整體而言， λ 可使變異數下降，但使估計誤差上升。在 λ 的選擇上，可透過赤池資訊評選準則 (Akaike information criterion, AIC)、貝氏資訊評選準則 (Bayesian information criterion, BIC) 或均方差 (mean square error, MSE) 等方式挑選，而實務上則大多由交叉驗證 (cross validation) 極小化 MSE 的方式所決定。

脊迴歸有助選出較有關聯的解釋變數，並使不重要解釋變數的係數接近於 0，降低資料的干擾雜訊 (noise)，藉此提供較為清晰的真實訊息。惟另一方面，脊迴歸僅能使不重要的解釋變數之係數接近於 0，而非等於 0，故模型仍將保留所有變數。

(二) LASSO

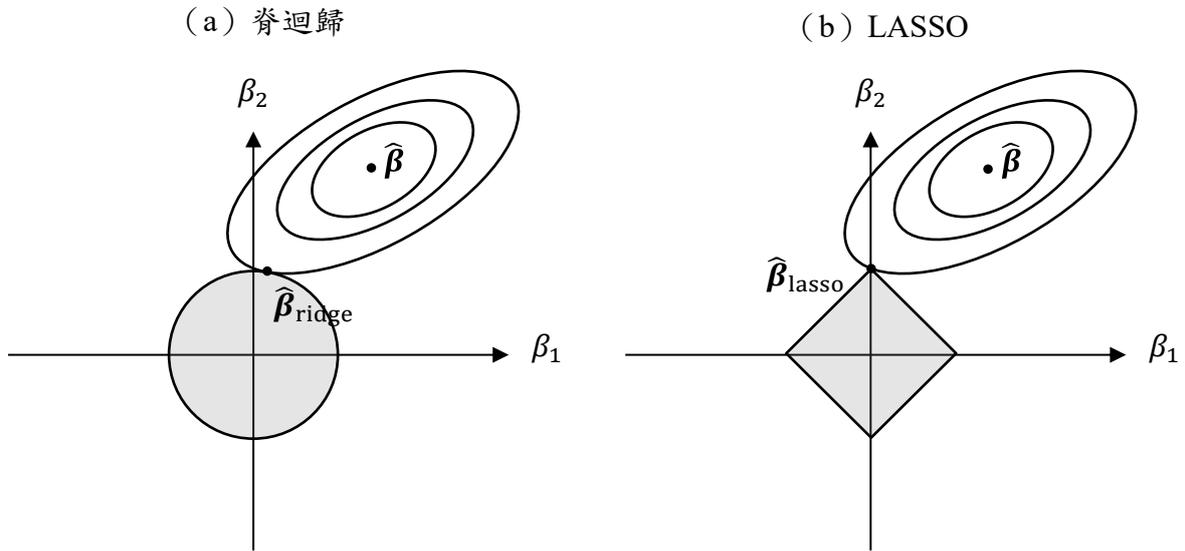
若在 OLS 損失函數加上懲罰項 $\lambda\|\beta\|$ ，則可得到 LASSO 的損失函數 (Tibshirani, 1996)：

$$\mathcal{L}_{\text{lasso}} = \|\mathbf{y} - \mathbf{X}\beta\|^2 + \lambda\|\beta\|. \quad (1.3)$$

其中， $\lambda\|\beta\|$ 亦稱為 L1 規範懲罰項 (L1 norm penalty term)，調控參數 λ 的概念與脊迴歸相同。LASSO 與脊迴歸最大的不同在於，脊迴歸僅能使不重要的解釋變數的係數接近於 0，但 LASSO 卻可使不重要的解釋變數的係數確實為 0。其原因在於，脊迴歸的懲罰項為平方和，但 LASSO 的懲罰項為絕對值總和，因此，脊迴歸的限制條件會形成圓形的可解區域，而 LASSO 的限制條件則會形成菱形的可解區域，進而得到不同的估計結果。

圖 1 為損失函數與限制條件的平面投影示意圖。圖中的橢圓區域為損失函數的平面投影、 $\hat{\beta}$ 為係數估計值，圖 1 (a)、(b) 中的圓形灰

圖 1：損失函數及限制條件示意圖



說明：橢圓區域、灰色區域分別為損失函數、懲罰項可解區域的平面投影。

資料來源：Tibshirani (1996)

色區域、菱形灰色區域則分別為脊迴歸、LASSO 懲罰項可解區域的平面投影，而損失函數與限制條件的交點即為估計結果。從圖中可知， β_2 為較重要之參數，圖 1 (a) 脊迴歸中的 β_1 估計值甚小但不為 0，而圖 1 (b) LASSO 中的 β_1 估計值則為 0。

由於 LASSO 可使不重要的解釋變數係數確實為 0，故相對脊迴歸，LASSO 較有助挑出較具影響力的變數或特徵 (feature)；但另一方面，若將較不重要的特徵全數移除，則可能會犧牲部分預測準確性。

(三) Elastic Net

由於脊迴歸與 LASSO 各有優缺，因此，Elastic Net 結合脊迴歸與 LASSO 兩者的懲罰項，冀能克服兩者的缺點 (Zou and Hastie, 2005)。Elastic Net 的損失函數為：

$$\mathcal{L}_{\text{ENet}} = \frac{\|\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\|^2}{2n} + \lambda \frac{1 - \alpha}{2} \|\hat{\boldsymbol{\beta}}\|^2 + \lambda \alpha \|\hat{\boldsymbol{\beta}}\|. \quad (1.4)$$

其中， λ 、 α 均為調控參數。 λ 的概念與脊迴歸及 LASSO 相同，而當 $\alpha = 0$ 時，Elastic Net 將回到脊迴歸的損失函數；當 $\alpha = 1$ 時，Elastic Net 則回到 LASSO 的損失函數。同樣地， λ 、 α 可透過 AIC、BIC 或 MSE 等方式挑選。Elastic Net 有助控制多重共線性的問題，使迴歸模型在解釋變數過多時 ($k > n$) 仍得以執行，同時亦可降低資料的干擾雜訊，得以在尋找較具影響力的變數時，兼顧預測的準確性。

在正規化迴歸的預測應用上，首先，Elastic Net 與前述之脊迴歸、LASSO 等方法皆屬線性模型，即假設解釋變數 (\mathbf{X}) 與目標變數 (\mathbf{y}) 之間為線性關係；因此，若欲探討變數之間的非線性關係，則應考慮非線性模型。其次，解釋變數的數值大小會影響正規化迴歸的求解過程，故原則上，應先將解釋變數予以標準化 (standardization) 後，再代入迴歸式求解。最後，正規化迴歸未含解釋變數本身的動態關係，故當欲預測的期數大於 1 時 ($h > 1$)，僅能以直接 (direct) 的方式進行預測，而無法以疊代 (iterative) 的方式進行預測。

二、貝氏 VAR

在總體經濟分析中，VAR 模型可捕捉不同變數當期與落後期之間的關係，據此估計各變數的衝擊反應、變異數分解等，並可用於預測，係廣受應用的多變量模型。一般而言，VAR 模型可表示如下：

$$\mathbf{y}_t = \mathbf{a}_0 + \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_p\mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(0, \boldsymbol{\Sigma}). \quad (2.1)$$

其中， \mathbf{y}_t 為 $M \times 1$ 的內生 (endogenous) 變數向量， \mathbf{a}_0 為 $M \times 1$ 的常數項向量， $\mathbf{A}_1, \dots, \mathbf{A}_p$ 為 $M \times M$ 的係數矩陣， p 為模型的落後期數。此外， $\boldsymbol{\varepsilon}_t$ 為 $M \times 1$ 的外生衝擊 (exogenous shock) 向量，服從平均數 0、變異數-共變異數矩陣為 $\boldsymbol{\Sigma}$ 之常態分配。

從 (2.1) 式可知，VAR 的參數數量共為 $M + p \times M^2$ 。由於待估參數眾多，變數增加將導致模型過度參數化 (over-parameterization) 或過度配適 (over-fitting) 等問題。在此情況下，透過貝氏方法賦予 VAR 模型某些先驗 (prior) 資訊，即可減緩計算上的限制，使得大量變數的大型 VAR 模型得以估計，此即貝氏 VAR 的概念 (Bańbura et al., 2010)。

至於先驗資訊的設定，一向係文獻上的重要議題。Giannone et al. (2015) 提出以資料為基礎的分層先驗 (hierarchical prior) 方法；亦即，將先驗資訊中的超參數 (hyperparameters) 視為待估計的額外參數，而超參數本身亦有其先驗資訊，此即超先驗 (hyperpriors)，可透過資料估計。藉由貝氏定理，其做法可說明如下²：

$$\begin{aligned} \mathbb{P}(\boldsymbol{\gamma}|\mathbf{y}) &\propto \mathbb{P}(\mathbf{y}|\boldsymbol{\gamma}, \boldsymbol{\theta})\mathbb{P}(\boldsymbol{\theta}|\boldsymbol{\gamma})\mathbb{P}(\boldsymbol{\gamma}), \\ \mathbb{P}(\mathbf{y}|\boldsymbol{\gamma}) &= \int \mathbb{P}(\mathbf{y}|\boldsymbol{\gamma}, \boldsymbol{\theta})\mathbb{P}(\boldsymbol{\theta}|\boldsymbol{\gamma}) d\boldsymbol{\theta}. \end{aligned}$$

其中， $\mathbf{y} = [\mathbf{y}_{p+1}, \dots, \mathbf{y}_T]'$ ， $\boldsymbol{\theta}$ 為 VAR 模型之係數及變異數等參數， $\boldsymbol{\gamma}$ 為超參數之集合。因此， $\mathbb{P}(\mathbf{y}|\boldsymbol{\gamma})$ 為資料 \mathbf{y} 在給定超參數 $\boldsymbol{\gamma}$ 下的邊際密度 (marginal density) 函數，即邊際概似 (marginal likelihood) 函數，可作為選擇超參數的準則；而在分層的做法中，可用於尋找超參數的後驗範圍。

在先驗分配的設定上，Normal–Wishart 分配為多變量常態分配的自然共軛³，許多常見的先驗分配皆來自此分配族 (distribution family)。令 $\mathbf{A} = [\mathbf{a}_0, \mathbf{A}_1, \dots, \mathbf{A}_p]'$ 且 $\boldsymbol{\beta} = \text{vec}(\mathbf{A})$ ，則 Normal–Wishart 之

² 此處有關 Giannone et al. (2015) 分層先驗方法的說明，主要取自 Kuschnig and Vashold (2021)。

³ 在共軛先驗下，可直接得到後驗分配的封閉形式 (closed form)，有助計算邊際概似值。

先驗可表示為

$$\boldsymbol{\beta}|\boldsymbol{\Sigma} \sim \mathcal{N}(\mathbf{b}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Omega}),$$

$$\boldsymbol{\Sigma} \sim \mathcal{IW}(\boldsymbol{\Psi}, \mathbf{d}).$$

其中， \mathbf{b} 、 $\boldsymbol{\Omega}$ 、 $\boldsymbol{\Psi}$ 及 \mathbf{d} 為超參數 $\boldsymbol{\gamma}$ 之低維向量函數（functions of a lower-dimensional vector）；而 $\mathcal{IW}(\cdot)$ 為「反 Wishart 分配」（inverse-Wishart distribution）。接著，Giannone et al.（2015）共考慮三種先驗資訊：

1. Minnesota 先驗（Minnesota prior）

此係假設模型中的變數皆具隨機漫步（random walk）性質（Litterman, 1980），廣為應用於總體實證之中。Minnesota 先驗可表示如下：

$$\mathbb{E}((\mathbf{A}_s)_{ij}|\boldsymbol{\Sigma}) = \begin{cases} 1, & i = j \text{ and } s = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (2.2)$$

$$\text{cov}((\mathbf{A}_s)_{ij}, (\mathbf{A}_r)_{kl}|\boldsymbol{\Sigma}) = \begin{cases} \lambda^2 \frac{1}{s^\alpha} \frac{\boldsymbol{\Sigma}_{ik}}{\psi_j / (d - M - 1)}, & l = j \text{ and } r = s, \\ 0, & \text{otherwise.} \end{cases} \quad (2.3)$$

其中， λ 、 α 與 ψ_j 為超參數，且各變數的 AR(1)係數為 1，其餘為 0。首先， λ 控制先驗資訊的緊度（tightness），若 $\lambda \rightarrow 0$ ，則先驗資訊最為重要，後驗與先驗分配完全相同；若 $\lambda \rightarrow \infty$ ，則資料提供的資訊最為重要，後驗分配等同於樣本估計結果⁴。其次，假設變數本身的落後期數愈多，則變異數下降愈快，而 α 即用於控制落後期變異數下降的速度。最後， ψ_j 為 $\boldsymbol{\Psi}$ 的第 j 個變數，用於控制其餘變數的落後期之先驗標準差。

2. 係數總和先驗（sum-of-coefficients prior）

⁴ 換句話說，「緊度」即代表指「先驗資訊」與「資料」之間的相對重要性。

在給定初始觀察值（initial observations）的條件下，VAR 模型可能隱含某些「決定性成分」（deterministic component）⁵。在一般未受限的 VAR 模型中，決定性成分通常具有「時間上的異質性」（temporal heterogeneity），樣本初始與末端的行為十分不同，可解釋在全樣本下，變數大部分的變動情形。因此，對此決定性成分予以某些額外的先驗限制，可能有助改善貝氏 VAR 的預測表現。Doan et al.（1984）的係數總和先驗即屬於此性質。

係數總和先驗的概念為「不變的預測值即為樣本起始時的最佳預測值」，其於資料中建構一組虛擬觀察值（dummy observations）：

$$\mathbf{y}_{M \times M}^+ = \text{diag}\left(\frac{\bar{\mathbf{y}}}{\mu}\right),$$

$$\mathbf{x}_{M \times (1+Mp)}^+ = [\mathbf{0}_{M \times 1}, \mathbf{y}^+, \dots, \mathbf{y}^+].$$

其中， $\bar{\mathbf{y}}$ 為 $M \times 1$ 的向量，由各變數的前 p 個觀察值的平均所組成。超參數 μ 控制變異數及先驗資訊的緊度，若 $\mu \rightarrow \infty$ ，則不含任何先驗資訊；若 $\mu \rightarrow 0$ ，則各方程式均有一個單根（unit root）且不具共整合（cointegration）關係⁶。

3. 虛擬初始觀察值先驗（dummy initial observation prior）

由於總體經濟的變數之間可能具有共整合關係，故 Sims（1993）將虛擬觀察值修改為

$$\mathbf{y}_{1 \times M}^{++} = \frac{\bar{\mathbf{y}}'}{\delta},$$

$$\mathbf{x}_{1 \times (1+Mp)}^{++} = \left[\frac{1}{\delta}, \mathbf{y}^{++}, \dots, \mathbf{y}^{++} \right].$$

⁵ 此係指「給定初始條件及 VAR 模型的估計係數下，對未來 y 的期望值」，即 $\mathbb{E}_p(\mathbf{y}_t | \mathbf{y}_1, \dots, \mathbf{y}_p, \hat{\boldsymbol{\beta}})$ 。

⁶ 故此時 VAR 模型中的單根數量與變數數量相同。

其中， $\bar{\mathbf{y}}$ 的定義同前。超參數 δ 控制先驗資訊的緊度，若 $\delta \rightarrow \infty$ ，則不含任何先驗資訊；若 $\delta \rightarrow 0$ ，則 VAR 中的變數將為其非條件均數（unconditional mean），模型含有不具飄移項（drift）的單根且數量未知，故可能存在共整合關係。

綜上所述，在 Giannone et al. (2015) 分層先驗的貝氏 VAR 中，須估計之超參數為 λ 、 μ 、 δ 與 ψ_j 。Giannone et al. (2015) 將 λ 、 μ 、 δ 的先驗分配設定為 Gamma 分配，而 $\psi_j/(d - M - 1)$ 則設為反 Gamma（inverse-Gamma）分配。

三、動態因子模型

因子模型係從大量的時間序列變數中，萃取相對少量且無法直接觀察到（unobserved）的潛在（latent）因子；而各變數非屬潛在因子的部分，則可代表某種衡量誤差（measurement error）或其獨自特性。因此，因子模型係將大量變數縮減至少量的共同因子，捕捉共同的波動走勢，藉此達到降低維度的目的。例如，國內生產毛額（Gross Domestic Product, GDP）可視作景氣循環再加上某些衡量誤差，股票報酬則可視作市場走勢再加上某些特定公司或產業的特性，而其中之「景氣循環」或「市場走勢」即可作為共同因子的概念。

一般而言，因子模型可表示為

$$\mathbf{X}_t = \mathbf{\Lambda}\mathbf{F}_t + \boldsymbol{\zeta}_t. \quad (3.1)$$

其中， \mathbf{X}_t 為 $M \times 1$ 的可觀察變數（observable variable）向量， \mathbf{F}_t 為 $N \times 1$ 的潛在因子向量， $\mathbf{\Lambda}$ 為 $M \times N$ 的因子負載（factor loadings）矩陣， $\boldsymbol{\zeta}_t$ 則為特徵干擾（idiosyncratic disturbances）向量或誤差項。在（3.1）式中， \mathbf{F}_t 的前後期並無自我相關，故僅為靜態（static）的形式。因此，

若欲考慮動態關係，則須假設 \mathbf{F}_t 具有自我迴歸性質：

$$\mathbf{F}_t = \mathbf{A}_t \mathbf{F}_{t-1} + \cdots + \mathbf{A}_p \mathbf{F}_{t-p} + \mathbf{e}_t. \quad (3.2)$$

其中， $\mathbf{A}_t, \dots, \mathbf{A}_p$ 為 $N \times N$ 的係數矩陣， \mathbf{e}_t 為誤差項。在考慮潛在因子的自我迴歸性質後，(3.1)式與(3.2)式即共同組成動態因子模型。

由此可知，動態因子模型具有狀態空間 (state space) 模型的形式；亦即，(3.1)式為測量方程式 (measurement equation)，描繪可觀察變數與不可直接觀察的潛在因子之間的關係，而(3.2)式則為移轉方程式 (transition equation)，用於描繪潛在因子之間的動態關係，其亦可視為 VAR(p)模型。例如，若為兩個因子($N = 2$)與落後兩期($p = 2$)的動態因子模型，則(3.2)式可表示為

$$\begin{bmatrix} f_{1,t} \\ f_{2,t} \\ f_{1,t-1} \\ f_{2,t-1} \end{bmatrix} = \begin{bmatrix} a_{1,1}^1 & a_{1,2}^1 & a_{1,1}^2 & a_{1,2}^2 \\ a_{2,1}^1 & a_{2,2}^1 & a_{2,1}^2 & a_{2,2}^2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} f_{1,t-1} \\ f_{2,t-1} \\ f_{1,t-2} \\ f_{2,t-2} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \mathbf{e}_t.$$

此外，在估計之前，須先決定因子的數量 N ，此一般可依據模型的配適程度 (goodness-of-fit) 或 Bai and Ng (2002) 的資訊準則決定。最後，即可透過 Kalman filter 或主成分分析 (principal components analysis, PCA) 等方法進行估計。由於變數數值大小會影響估計結果，故變數應予以標準化。

在預測時，可透過目標變數與潛在因子之間的關係建構預測模型。值得提醒的是，當欲預測的期數大於1時 ($h > 1$)，除了直接預測的方式之外，由於(3.2)式屬於 VAR(p)的形式，其描繪 \mathbf{F}_t 當期與落後期之間的動態關係，故亦可透過 VAR(p)以疊代的方式進行預測。更多有關動態因子模型之介紹，可參見 Stock and Watson (2011)。

參、經濟預測的相關議題

在經濟預測上，除了模型建構之外，尚有許多重要議題。本節依序簡介「預測組合」與「結構改變」等兩項。首先，無論何種預測模型，皆可能存在某些設定上的錯誤 (misspecification)，故若能將不同的預測值以某種方式組合，或可提升預測的準確性，此即預測組合的概念。其次，一般在預測時，通常假設模型可刻劃隱含的資料生成過程 (data generating process, DGP)，且在預測期間內，經濟結構無甚改變。但實際上，經濟體系常有無法預料的變遷，使得模型通常會不同於 DGP，且 DGP 亦可能具有非定態 (non-stationary) 性質，皆會使模型的預測表現不甚理想。因此，如何處理結構改變，亦為文獻上的重要議題。

一、預測組合

如前所述，各種模型皆可能存在設定錯誤，故在不同的時期或經濟環境下，單一最佳模型往往甚難尋求；而預測組合即綜合不同預測值的資訊，進而達到降低預測誤差的目的。在點預測 (point forecast) 的應用上，一般常見的組合方式包括簡單平均 (simple average)、削減平均 (trimmed mean)、以 MSE 的倒數 (inverse MSE) 作為權重、以 MSE 排序的倒數 (inverse rank) 作為權重或 OLS 迴歸等幾種方式。

1. 簡單平均：對各個不同的預測值取簡單平均，可表示為

$$f_{ew} = \frac{1}{P} \sum_{i=1}^P f_i.$$

其中， f_i 為不同的預測值， P 為預測值的數量。由於此係直接給予

各預測值相同的權重，故做法簡單且不具估計誤差，缺點則為容易受離群值 (outlier) 所影響。但文獻上發現，簡單平均的預測表現往往最佳，更勝於其他較為複雜的組合方法，此即「預測組合謎團」(forecast combination puzzle) (Stock and Watson, 2004)⁷。

2. 削減平均：剔除較高與較低的部分預測值後再取平均，以避免離群值之影響。可表示為

$$f_{tm} = \frac{1}{P(1-2\lambda)} \sum_{i=\lambda P+1}^{(1-\lambda)P} f_i.$$

其中， λ 為較高與較低的預測值之占比。例如，若 $\lambda = 0.1$ ，則代表剔除最高與最低的 10% 的預測值後，再對中間剩餘 80% 的預測值取平均。

3. 以 MSE 的倒數作為權重：若模型的 MSE 愈低，則表示模型愈準確，故可給予愈高的權重。可表示為

$$f_{i-mse} = \frac{MSE_i^{-1}}{\sum_{i=1}^P MSE_i^{-1}} \sum_{i=1}^P f_i.$$

其中， MSE_i^{-1} 即為第 i 個預測模型之 MSE 的倒數。

4. 以 MSE 排序的倒數作為權重：同樣地，此做法亦給予較準確的模型較高的權重，但其權重並非直接依據 MSE 計算，而係將各模型依據 MSE 之數值大小進行排序後，再依據其排序的倒數給予權重。可表示為

$$f_{i-rank} = \frac{Rank_i^{-1}}{\sum_{i=1}^P Rank_i^{-1}} \sum_{i=1}^P f_i.$$

⁷ 有關預測組合謎團之解釋與討論，可參見 Claeskens et al. (2016) 或 Chan and Pauwels (2018)。

其中，Rank即為 MSE 的排序，根據 MSE 的數值，由小至大依序為1,2,...,P。

5. OLS 迴歸：利用 OLS 估計各預測值與目標變數之間的關係，再利用估計係數作為權重，將各預測值組合為新的預測值。亦即，須先估計

$$y = \alpha + \sum_{i=1}^P w_i f_i + \varepsilon,$$

其中， y 為目標變數， ε 為迴歸殘差， α 、 w_i 為待估參數。在估得各預測值的權重 \hat{w}_i 及常數項 $\hat{\alpha}$ 後，即可得到其組合預測值

$$f_{ols} = \hat{\alpha} + \sum_{i=1}^P \hat{w}_i f_i.$$

由於常數項 $\hat{\alpha}$ 存在，故即便各預測值均有所偏誤 (biased)，此做法仍可得到不偏 (unbiased) 的組合預測值；另一方面，由於 \hat{w}_i 均為係數估計值，故其意涵可能不易解讀。

預測組合的議題由來已久，且仍持續發展未艾，此處僅簡述幾種較為基本的預測組合方法。有關預測組合更完整的介紹，可參見 Timmermann (2006) 之探討⁸。

二、結構改變

結構改變係指參數發生不規則 (irregular)、間斷 (discrete) 且永久性 (permanent) 的改變 (Eklund et al., 2013)。例如，變數之間的關係可能隨時間變化，進而造成模型參數改變；或是資料或變數本身具

⁸ 除了點預測之外，密度預測 (density forecast) 亦可予以組合，可參見 Hall and Mitchell (2007) 或 Kascha and Ravazzolo (2010) 之討論。

有單根或長記憶 (long memory) 行為，使得資料會隨時間而不穩定。當結構穩定時，可透過最佳的樣本內模型刻劃 DGP，再藉由即時資料進行預測；此時，樣本外的預測表現應與樣本內的配適結果相似。然而，若結構發生改變，則可能造成迴歸係數或誤差變異數產生變動，進而影響模型的預測結果。由於使用模型係基於模型結構固定之前提，故在結構改變可能存在的情況下，使用最佳樣本內模型的做法或非最適，而應再行考量。

首先，應用滾動視窗 (rolling window) 進行估計，或是參數因時而變 (time-varying)、馬可夫轉換 (Markov-switching) 等模型，皆為文獻上常見的做法。前兩者較適用於平滑或漸進式的結構轉變，而後者則可用於較屬突發 (adrupt) 性質的狀態轉換。

其次，認定結構改變的時點，並僅以該時點後的樣本進行估計，亦為可行的做法之一。一般而言，若結構改變的時點已知，則可借助 Chow 檢定進行確認 (Chow, 1960)⁹；若結構改變的時點未知，則可借助 Bai-Perron 檢定 (Bai and Perron, 1998, 2003)。惟整體而言，結構改變之認定仍有許多限制。例如，較小的改變不易發現、檢定方法不適用於較頻繁的改變、檢定方法多需保留 5–15% 的端點樣本、近期的改變難以即時得知等¹⁰。再者，若結構改變的時點較為近期，則結構改變後的樣本觀察值可能不多，亦將增加估計結果的不確定性。Pesaran and Timmermann (2005) 建議，用於估計預測模型的樣本觀察值，數量至少應為預測模型參數數量的二至三倍。

⁹ 此外，對於某些已知的結構改變，亦可透過模型予以刻劃，或是以虛擬變數捕捉，如 Lenza and Primiceri (2022) 或 Cascaldi-Garcia (2022) 等文獻對 COVID-19 疫情的處理方法。

¹⁰ Chu et al. (1996) 曾基於「波動檢定」(fluctuation test) 提出即時監控結構改變的做法，可參見 Zeileis et al. (2002) 的說明。

另一方面，部分文獻則提出較為「穩健」(robust)的做法；亦即，在進行預測的過程中，不另行考量有關結構改變的資訊。例如，若結構改變源自 VAR 模型中的決定性趨勢，則可透過「截距校正」(intercept correction)或「過度差分」(over-differencing)的方式改善預測績效 (Clements and Hendry, 1999)。前者追蹤模型過去的預測績效，若模型有系統性的高估或低估，則透過截距校正其偏誤¹¹；後者則透過二階差分消除模型中的決定性趨勢。除此之外，亦可混合 (pool) 不同的模型或資訊，藉此得到穩健的預測值。例如，預料外的變遷對各模型皆可能有不同的影響，故若組合不同模型的預測值，則可能改善預測績效 (Clements and Hendry, 2004)；在結構改變已知的情況下，透過不同長度滾動視窗建構預測值，再予以結合，亦可提供穩健的預測表現 (Pesaran and Timmermann, 2007)；而混合資訊的做法則包括擴散指標 (diffusion index)、動態因子模型，或是使用混合頻率 (mixed-frequency) 的資料等。

另有文獻建議，可適度納入結構改變前的樣本觀察值。Kapetanios et al. (2015) 歸納三種做法：(1) 基於最近的預測表現，決定最適的滾動視窗長度；(2) 以指數加權移動平均 (exponentially weighted moving average) 方法，賦予愈近期的資料愈大的權重、愈早期的資料愈低的權重；(3) 以「狀態」而非「時間遠近」決定權重。例如，當經濟處於復甦階段，則賦予過去經濟復甦時期的資料較大的權重、經濟衰退時期的資料較低的權重，可能有較佳的預測結果。

¹¹ 此即文獻上之截距項或常數項調整值 (add factor)。

肆、應用範例：預測台灣的經濟成長率

中央銀行在制定貨幣或金融穩定政策時，未來經濟展望一向為重要參考依據。然而，近年經濟金融情勢波動加劇，不確定性增加，以致經濟預測的難度更甚以往。由於善用大量變數提供的資訊，可能對經濟預測有所助益，因此，本節應用前述介紹之內容，就台灣經濟成長率之預測進行實證分析，以作為說明範例。

一、簡介

近年來，「大數據」(big data) 議題興起；而事實上，已有許多國內文獻致力於大量變數之應用，並發現大量變數有助提升經濟預測的準確性。例如，徐士勛等 (2005) 應用因子模型從大量總體變數萃取資訊，並發現其對經濟成長率的預測表現甚佳；張天惠 (2012)、張天惠與朱浩榜 (2020) 應用動態因子模型及大量金融變數建構台灣的金融情勢指數 (financial conditions index, FCI)，並發現 FCI 領先實質面的經濟變數，且可提供有助經濟預測的額外資訊；管中閔等 (2014) 亦利用眾多金融變數建構台灣的 FCI，且其 FCI 亦具領先大多數實質面總體變數的特質。

近來，吳俊毅與朱浩榜 (2020) 透過混合頻率資料抽樣 (mixed-data sampling, MIDAS) 模型及大量月頻經濟指標，發現月頻之即時資訊確實有助當季經濟成長率之即時預報 (nowcast)，且結合因子模型之預報績效相對較佳；蕭宇翔與林依伶 (2020) 則將大量總體變數運用於景氣狀態之預測，並發現機器學習方法對未來 1 至 6 個月的預測表現普遍較佳，而因子模型對未來 1 至 3 個月亦有不錯的預測表現。

由上可知，大量變數的資訊確實有助台灣的總體經濟預測；而在預測方法上，目前國內文獻大多聚焦在因子模型的應用。爰此，除了

動態因子模型之外，本報告亦考量正規化迴歸與貝氏 VAR 等模型，並進一步考量預測組合，藉此探討不同的大量變數模型及預測組合方法，對於提升預測績效的能力。

二、資料說明

首先，參照文獻上的做法（如 Stock and Watson, 2004），定義未來 h 期的年化後（annualized）經濟成長率為

$$y_{t+h}^h = 100 \times \left(\frac{4}{h}\right) \ln \left[\frac{\text{GDP}_{t+h}}{\text{GDP}_t} \right].$$

此即欲預測的目標變數，代表在 t 時點上，未來 h 期 GDP 的平均變動情形。而各期的經濟成長率則為

$$y_t = 100 \times \left(\frac{4}{1}\right) \ln \left[\frac{\text{GDP}_t}{\text{GDP}_{t-1}} \right].$$

即在 t 時點上，年化後的 GDP 季變動率。

除了 GDP 之外，其餘變數包括工業生產指數、實質出口總值、GDP 平減指數、消費者物價指數、失業率、每人每月總薪資等經濟活動方面的變數，以及新台幣對美元匯率、新台幣實質有效匯率指數、金融業隔夜拆款利率、10 年期中央政府公債次級市場利率、長短期利差、股票市場加權指數、股票市場總市值、股票市場成交金額、貨幣總計數 M1B、M2、全體貨幣機構放款與投資等金融方面的變數¹²。資料來源詳見表 1，樣本期間涵蓋 1996 年第 1 季至 2022 年第 4 季¹³。

¹² 其中，實質出口總值係以「新台幣計價之出口物價指數」平減「新台幣計價之出口總值」計算；長短期利差則以「10 年期公債次級市場利率」減去「金融業隔夜拆款利率」計算。

¹³ 由於本報告僅作為應用範例，相較於過去國內文獻，使用的變數數量相對較少。

表 1：實證變數說明

變數	資料來源	處理方式
經濟活動面		
GDP	主計總處	變動率
工業生產指數	經濟部統計處	變動率
實質出口總值	財政部統計處、主計總處 ^a	變動率
GDP 平減指數	主計總處	變動率
消費者物價指數	主計總處	變動率
失業率	主計總處	差分
每人每月總薪資	主計總處	變動率
金融面		
新台幣對美元匯率	中央銀行	變動率
新台幣實質有效匯率	國際清算銀行	變動率
金融業隔夜拆款利率	中央銀行	差分
10 年期中央政府公債次級市場利率	中央銀行	差分
長短期利差	中央銀行 ^b	差分
貨幣總計數 M1B	中央銀行	變動率
貨幣總計數 M2	中央銀行	變動率
全體貨幣機構放款與投資	中央銀行	變動率
股票市場加權指數	台灣證券交易所	變動率
股票市場總市值	台灣證券交易所	變動率
股票市場成交金額	台灣證券交易所	變動率

說明：^a 實質出口總值係以「新台幣計價之出口物價指數」平減「新台幣計價之出口總值」計算，其中，出口總值資料來源為財政部統計處、出口物價指數資料來源為主計總處；^b 長短期利差係以「10 年期公債次級市場利率」減去「金融業隔夜拆款利率」計算，兩者資料來源皆為中央銀行。

由於 GDP 為季頻資料，故在實證分析之前，須先將各變數均轉換為季頻資料；再者，經濟活動易受季節性因素影響，故再以 X-13ARIMA-SEATS 方法對 GDP 等經濟活動方面的變數進行季節調整。接著，為了避免變數有趨勢或單根的情形，因此以變動率的方式進行變數處理。變動率的計算方式與經濟成長率相同：

$$100 \times \left(\frac{4}{1}\right) \ln \left[\frac{X_t}{X_{t-1}} \right].$$

其中， X_t 為各變數在 t 時點上的原始值。至於失業率、利率及利差等單位已為百分比或百分點之變數，則改以差分的方式進行處理：

$$X_t - X_{t-1}.$$

同樣地， X_t 為變數在 t 時點上的原始值。關於各變數的處理方式，亦可參見表 1。

三、預測方法

在預測方法上，首先，設定 AR 模型作為比較的基準 (benchmark) 模型：

$$y_{t+h}^h = \alpha_c + \sum_{i=0}^4 \alpha_i y_{t-i} + \varepsilon_{t+h}.$$

其中， y_{t+h}^h 、 y_{t-i} 如前述定義，而解釋變數則考量當期至落後 4 期的經濟成長率，以納入過去一年內的資訊。由於 AR 模型僅含有變數本身的資訊，故可藉此檢驗其餘變數或方法是否可提供額外的預測資訊。在預測方面，AR 模型可透過直接與疊代等兩種方式得到預測值，因此皆予考量。

其次，脊迴歸、LASSO 與 Elastic Net 等三種正規化迴歸的模型設定分別如 (1.2) 式、(1.3) 式與 (1.4) 式所示。在模型的解釋變數中，不僅考量各變數的當期值，亦考量各變數的落後 1 至 4 期，且變數亦包括 GDP 本身。至於迴歸式的調控參數，則透過交叉驗證極小化 MSE 的方式決定。另在正規化迴歸的預測方面，如前所述，僅能透過直接預測的方式得到預測值。

接著，貝氏 VAR 以 Giannone et al. (2015) 的分層先驗方法估計。其先驗分配為 (2.2) 式、(2.3) 式之 Minnesota 先驗，惟各變數已經過變動率或差分處理，故先驗平均設為 0；在超參數方面則依據

Giannone et al. (2015)，設定 λ 為眾數 0.2 的 Gamma 分配、 ψ_j 為尺度參數 (scale parameter) 與型態參數 (shape parameter) 皆為 $(0.02)^2$ 的反 Gamma 分配， α 則設定為 2。至於「係數總和」與「虛擬觀察值」先驗，由於變數已取變動率或差分，應無單根或共整合之情形，故未予考量。在模型估計上，採用「馬可夫鏈蒙地卡羅」(Markov chain Monte Carlo) 方法計算，抽樣次數為 15,000 次，且為了避免初始值影響估計結果，前 5,000 筆抽樣結果予以捨棄 (burn-in)。最後在預測上，此處以疊代的方式得到預測值。

最後，動態因子模型設定如 (3.1) 式與 (3.2) 式。由於此處採用的實證變數涵蓋經濟活動與金融等兩個面向 (見表 1)，故因子數量設定為 2；而因子模型的落後期則參照其他模型，納入過去一年的資訊，故設定為 4。在估計方法上，採用 Doz et al. (2012)「類最大概似」(quasi-maximum likelihood, quasi-ML) 之方法估計。在估得各變數的共同因子後，即可建構預測模型：

$$y_{t+h}^h = \rho_c + \sum_{i=0}^4 \rho_i y_{t-i} + \sum_{r=1}^2 \phi_r f_{r,t} + \epsilon_{t+h}.$$

其中， $f_{r,t}$ 即為共同因子，其餘變數的定義同前。而此處係以直接預測的方式得到預測值。

綜上所述，在大量變數的預測模型方面，共可得到脊迴歸、LASSO、Elastic Net、貝氏 VAR 與動態因子模型等五個預測值。因此，若再考量簡單平均、20%的削減平均等兩種預測組合方法，則可得到七個不同的預測值。以下就此七種方法與兩種 AR 模型預測值 (直接及疊代預測) 的預測績效進行比較。

為評估模型的預測績效，須以樣本外的方式進行預測。惟長期以

來，台灣的經濟結構發生許多變遷，故以滾動迴歸 (rolling regression) 的方式進行樣本外預測，而滾動視窗選擇為 60 季。預測期間為未來一季 ($h = 1$)、未來半年 ($h = 2$)、未來一年 ($h = 4$)、未來一年半 ($h = 6$) 及未來兩年 ($h = 8$)。資料期間涵蓋 1996 年第 1 季至 2022 年第 4 季，而在考量變數處理、模型落後期、視窗長度及預測期間後，以下對 2014 年第 1 季至 2020 年第 4 季的未來經濟成長率進行樣本外預測。最後，在得到各預測方法的預測值之後，即可計算不同預測期間下，各預測值與實際值之間的均方根差 (root-mean-square error, RMSE) 並比較其大小，藉此評估各模型的預測績效。

四、實證結果

表 2 列示基準模型與各預測方法之相對 RMSE，基準模型為 AR 模型。表中各列為不同的大量變數模型與預測組合方法、各欄為不同的預測期間，而相對 RMSE 則以各預測方法之 RMSE 除以基準模型之 RMSE，故當相對 RMSE 小於 1 時，各預測方法的預測表現較佳，顯示大量變數的預測模型或預測組合方法有助提升預測準確度。另在表 2 (a)、(b) 中，基準模型的預測值分別為 AR 模型以疊代、直接預測等兩種方式得到的預測值。此外，表中亦列出 AR 模型之 RMSE 供參。

表 2 的呈現方式說明如下。例如，在表 2 (a) 中，當 $h = 1$ 時，AR 模型以疊代方式預測的 RMSE 為 3.830%，且脊迴歸與 AR 模型的相對 RMSE 為 1.009，顯示脊迴歸的預測績效較差，無法改善預測的準確度；而 LASSO 與 AR 模型的相對 RMSE 則為 0.970，顯示 LASSO 的預測績效較佳，其可從大量變數中取得有關未來 1 期經濟成長的資訊。

表 2：各預測方法與 AR 模型的相對 RMSE

	預測期間				
	$h = 1$	$h = 2$	$h = 4$	$h = 6$	$h = 8$
(a) 基準模型：AR (疊代預測)					
脊迴歸	1.009	1.022	0.935	0.920	0.944
LASSO	0.970	0.918	1.111	1.021	0.965
Elastic Net	0.975	0.928	1.048	0.897	0.996
貝氏 VAR	1.405	1.349	1.140	1.316	1.124
動態因子	1.020	1.086	0.979	1.022	1.022
簡單平均	0.987	0.989	0.972	0.902	0.921
削減平均	0.961	0.959	0.973	0.900	0.926
RMSE (%)	3.830	2.914	2.272	1.766	1.516
(b) 基準模型：AR (直接預測)					
脊迴歸	0.948	0.942	0.884	0.841	0.891
LASSO	0.911	0.846	1.050	0.933	0.911
Elastic Net	0.916	0.855	0.990	0.820	0.940
貝氏 VAR	1.320	1.243	1.078	1.203	1.062
動態因子	0.958	1.000	0.925	0.934	0.965
簡單平均	0.927	0.911	0.918	0.825	0.870
削減平均	0.903	0.883	0.920	0.823	0.875
RMSE (%)	4.077	3.163	2.404	1.931	1.606

- 說明：1. 表中各列為不同的預測方法，表中數值為相對 RMSE，計算方式為不同預測方法之 RMSE 除以基準模型之 RMSE，小於 1 標示為粗體，代表預測方法之 RMSE 較小，預測績效較佳。
 2. 表(a)、(b)中的基準模型分別為 AR 模型以疊代、直接預測方式得到的預測值。
 3. 表中各預測方法均採滾動方式進行樣本外預測，預測目標為 2014 年第 1 季至 2020 年第 4 季未來 h 期之經濟成長率 y_{t+h}^h 。
 4. 簡單平均為脊迴歸、LASSO、Elastic Net、貝氏 VAR 與動態因子等模型預測值之簡單平均；削減平均則為其 20% 的削減平均。

從表 2 (a) 可知，首先，當 AR 模型以疊代方式進行預測時，三種正規化迴歸的預測表現可在多數預測期間下略勝一籌，動態因子模型僅在 $h = 4$ 時表現略佳，而貝氏 VAR 則表現稍差。其次，若考量上述五種模型的預測組合，則可發現無論在何種預測期間下，兩種預測組合方法的表現均勝於 AR 模型。因此，若能有效綜合不同預測值的

資訊，則有助達成降低預測誤差的目的。

接著，從表 2 (b) 可知，AR 模型在直接預測時的 RMSE 在各期間下均高於疊代預測，顯示直接預測的表現稍差，並使各預測方法之相對 RMSE 均呈下降。因此，在三種正規化迴歸中，僅 LASSO 在 $h = 4$ 時的 RMSE 略高，動態因子模型的表現亦普遍轉佳，僅貝氏 VAR 仍表現略差。另就預測組合而言，由於兩種組合方法的 RMSE 原本即低於以疊代方式進行預測的 AR 模型，故在 AR 模型改以直接方式進行預測後，相對 RMSE 均進一步降低。

綜上所述，應用大量變數的預測模型，確實可提供有關未來經濟成長的額外資訊；若能再綜合不同預測值的資訊，則有助進一步降低誤差，提升預測的準確度。值得注意的是，本報告作為應用範例，使用的變數數量相對較少；若能參照過去國內文獻使用更多變數，或可進一步提供更有效的預測資訊¹⁴。再者，此處僅採用較為基本的模型設定，而近來大量變數模型的發展甚快且應用頗多，故模型運用亦有進一步延伸的空間。可資參考的相關研究頗豐，例如，Kotchoni et al. (2019) 應用大量變數模型於總體經濟預測，而 Medeiros et al. (2021) 則應用機器學習方法於通膨預測。此外，亦可參見 Masini et al. (2023) 有關機器學習與時間序列預測的討論。

¹⁴ 例如，徐士勛等 (2005)、吳俊毅與朱浩榜 (2020)、蕭宇翔與林依伶 (2020) 分別採用 81、79、82 個總體變數 (含金融變數在內)；而在 FCI 的編製上，張天惠 (2012)、管中閔等 (2014)、張天惠與朱浩榜 (2020) 則分別採用 12、19、46 個金融變數。

伍、結論與建議

近年由於 COVID-19 疫情衝擊、通膨驟升、俄烏戰爭、全球央行進入升息循環、歐美銀行業危機等因素，全球經濟金融環境快速變遷，不確定性大於以往，使得經濟預測的難度上升，致央行貨幣或金融穩定政策迎來挑戰。在快速變遷的環境下，本次研習講授之模型與議題切合時宜，頗有助益。

首先，大量變數或可提供較多資訊；但若欲同時納入大量變數，則參數過多將影響估計及預測結果，故可透過縮減參數、先驗資訊或萃取共同因子等方式降低維度，進而得到有用的預測模型。本報告應用正規化迴歸（包括脊迴歸、LASSO 與 Elastic Net）、貝氏 VAR 及動態因子模型等方法，發現大量變數確實可提供有關未來經濟成長的額外資訊。其次，由於各模型皆可能存有設定錯誤的問題，故若能組合不同模型的預測值，則準確性或可提升。而本報告的實證結果亦顯示，若能有效綜合不同預測值的資訊，確實有助進一步降低預測誤差。

接著，經濟體系常有無法料及的變遷，導致模型的預測表現不甚理想，故文獻上提出許多結構改變的處理方式，可資未來研究參考。另為傳達經濟預測的不確定性，許多央行已嘗試公布不同形式的密度預測，助其與民眾溝通，俾利貨幣與金融穩定政策之施行。

在本次研習課程中，講師除講述相關課程外，另於課程期間安排 EViews、R 與 MATLAB 等應用軟體之實務操作，對理論學習及未來實際運用均助益良多。此外，課程亦邀請尼泊爾央行（Nepal Rastra Bank）的官員，分享該國的經濟情勢、預測模型及預測工作之相關經驗，故亦可藉此機會與其他國家學員交流討論。因此，參與本次課程，當有助增進預測分析及研究能力。

參考資料

- 吳俊毅、朱浩榜 (2020), 「即時預報台灣的經濟成長率：MIDAS 模型之應用」, 《中央銀行季刊》, 42(1), 59–84。
- 徐士勛、管中閔、羅雅慧 (2005), 「以擴散指標為基礎之總體經濟預測」, 《臺灣經濟預測與政策》, 36(1), 1–28。
- 張天惠 (2012), 「我國金融情勢指數與總體經濟預測」, 《中央銀行季刊》, 34(2), 11–41。
- 張天惠、朱浩榜 (2020), 「台灣金融情勢與經濟預測」, 《經濟論文叢刊》, 48(1), 77–106。
- 管中閔、徐之強、黃裕烈、徐士勛 (2014), 「台灣金融情勢指數與總體經濟關係」, 《臺灣經濟預測與政策》, 44(2), 103–132。
- 蕭宇翔、林依伶 (2020), 「臺灣景氣狀態之預測」, 《臺灣經濟預測與政策》, 51(1), 1–56。
- Bai, J., and S. Ng (2002), “Determining the Number of Factors in Approximate Factor Models,” *Econometrica*, 70(1), 191–221.
- Bai, J., and P. Perron (1998), “Estimating and Testing Linear Models with Multiple Structural Changes,” *Econometrica*, 66(1), 47–78.
- Bai, J., and P. Perron (2003), “Computation and Analysis of Multiple Structural Change Models,” *Journal of Applied Econometrics*, 18(1), 1–22
- Bañbura, M., D. Giannone, and L. Reichlin (2010), “Large Bayesian Vector Auto Regressions,” *Journal of Applied Econometrics*, 25(1), 71–92.
- Cascaldi-Garcia, D. (2022), “Pandemic Priors,” International Finance Discussion Papers No.1352, Board of Governors of the Federal Reserve System.
- Claeskens, G., J. R. Magnus, A. L. Vasnev, and W. Wang (2016), “The Forecast Combination Puzzle: A Simple Theoretical

- Explanation,” *International Journal of Forecasting*, 32(3), 754–762.
- Doz, C., D. Giannone, and L. Reichlin (2012), “A Quasi-Maximum Likelihood Approach for Large, Approximate Dynamic Factor Models,” *Review of Economics and Statistics*, 94(4), 1014–1024.
- Clements, M. P., and D. F. Hendry (1999), *Forecasting Non-Stationary Economic Time Series*, Cambridge, Massachusetts, MIT.
- Clements, M. P., and D. F. Hendry (2004), “Pooling of Forecasts,” *Econometrics Journal*, 7(1), 1–31.
- Chan, F., and L. L. Pauwels (2018), “Some Theoretical Results on Forecast Combinations,” *International Journal of Forecasting*, 34(1), 64–74.
- Chow, G. (1960), “Tests of Equality between Sets of Coefficients in Two Linear Regressions,” *Econometrica*, 28(3), 591–603.
- Chu, C.-S. J., M. Stinchcombe, and H. White (1996), “Monitoring Structural Change,” *Econometrica*, 64(5), 1045–1065.
- Doan, T., R. Litterman, and C. Sims (1984), “Forecasting and Conditional Projection Using Realistic Prior Distributions.” *Econometric Reviews*, 3(1), 1–10.
- Eklund, J., G. Kapetanios, and S. Price (2013), “Robust Forecast Methods and Monitoring during Structural Change” *Manchester School*, 81(53), 3–27.
- Giannone, D., M. Lenza, and G. E. Primiceri (2015), “Prior Selection for Vector Autoregressions,” *Review of Economics and Statistics*, 97(2), 436–451.
- Hall, S. G., and J. Mitchell (2007), “Combining Density Forecasts,” *International Journal of Forecasting*, 23(1), 1–13.
- Hoerl, A. E., and R. W. Kennard (1970), “Ridge Regression: Biased Estimation for Nonorthogonal Problems,” *Technometrics*, 12(1), 55–67.

- Kapetanios, G., S. Price, and S. Stone (2015), “Forecasting GDP in the Presence of Breaks: When Is the Past a Good Guide to the Future?” *Bank Underground*, 20 August.
- Kascha, C., and F. Ravazzolo (2010), “Combining Inflation Density Forecasts,” *Journal of Forecasting*, 29(1–2), 231–250.
- Kotchoni, R., M. Leroux, and D. Stevanovic (2019), “Macroeconomic Forecast Accuracy in a Data-Rich Environment,” *Journal of Applied Econometrics*, 34(7), 1050–1072.
- Kuschnig, N., and L. Vashold (2021), “BVAR: Bayesian Vector Autoregressions with Hierarchical Prior Selection in R,” *Journal of Statistical Software*, 100(14), 1–27.
- Lenza, M., and G. E. Primiceri (2022), “How to Estimate a VAR after March 2020,” *Journal of Applied Econometrics*, 37(4), 688–699.
- Litterman, R. B. (1980), “A Bayesian Procedure for Forecasting with Vector Autoregressions,” *Journal of Business and Economic Statistics*, 4(1), 25–38.
- Masini, R. P., M. C. Medeiros, and E. F. Mendes (2023), “Machine Learning Advances for Time Series Forecasting,” *Journal of Economic Surveys*, 37(1), 76–111.
- Medeiros, M. C., G. F. Vasconcelos, Á. Veiga, and E. Zilberman (2021), “Forecasting Inflation in a Data-Rich Environment: the Benefits of Machine Learning Methods,” *Journal of Business and Economic Statistics*, 39(1), 98–119.
- Pesaran, M. H., and A. Timmermann (2005), “Small Sample Properties of Forecasts from Autoregressive Models under Structural Breaks,” *Journal of Econometrics*, 129(1–2), 183–217.
- Pesaran, M. H., and A. Timmermann (2007), “Selection of Estimation

- Window in the Presence of Breaks,” *Journal of Econometrics*, 137(1), 134–161.
- SEACEN (2023), 「Forecasting for Monetary and Financial Stability Policies」研習課程講義。
- Sims, C. A. (1993), “A Nine-Variable Probabilistic Macroeconomic Forecasting Model,” in: J. H. Stock, and M. W. Watson, (eds.), *Business Cycles, Indicators and Forecasting*, 179–212, University of Chicago Press.
- Stock, J. H., and M. W. Watson (2004), “Combination Forecasts of Output Growth in a Seven-Country Data Set,” *Journal of Forecasting*, 23, 405–430.
- Stock, J. H., and M. W. Watson (2011), “Dynamic Factor Models,” in: M. P. Clements, and D. F. Hendry, (eds.), *Oxford Handbook of Economic Forecasting*, 35–60, Oxford University Press.
- Tibshirani, R. (1996), “Regression Shrinkage and Selection via the LASSO,” *Journal of the Royal Statistical Society: Series B*, 58(1), 267–288.
- Timmermann, A. (2006), “Forecast Combinations,” in: G. Elliott, C. Granger, and A. Timmermann, (eds.), *Handbook of Economic Forecasting*, Vol.1, 135–196, Elsevier, Amsterdam.
- Zeileis, A., F. Leisch, K. Hornik, and C. Kleiber (2002), “Strucchange: An R Package for Testing for Structural Change in Linear Regression Models,” *Journal of Statistical Software*, 7(2), 1–38.
- Zou, H., and T. Hastie (2005), “Regularization and Variable Selection via the Elastic Net,” *Journal of the Royal Statistical Society: Series B*, 67(2), 301–320.