



Phép lọc tuyến tính và vấn đề khử xu hướng của chuỗi thời gian: Nghiên cứu thực nghiệm tại Việt Nam

BÙI THỊ THIÊN MỸ^{a,*}, NGUYỄN THỊ YẾN^a

^a Trường Đại học Ngân hàng TP. Hồ Chí Minh

THÔNG TIN	TÓM TẮT
Ngày nhận: 29/06/2021 Ngày nhận lại: 21/02/2022 Duyệt đăng: 28/02/2022 Mã phân loại JEL: C19; C22 Từ khóa: Phép lọc Baxter-King; Phép lọc Hodrick- Prescott; Phép lọc tuyến tính. Keywords: Baxter-King filter; Hodrick-Prescott filter; Linear filters.	Phép lọc là công cụ phổ biến để phân tích một chuỗi thời gian thành các thành phần xu hướng và chu kỳ. Tính chất chung của các phép lọc là giữ lại một số thành phần trong chuỗi thời gian gốc, đồng thời làm ảnh hưởng đến biên độ dao động và pha của chuỗi dữ liệu nhận được sau phép lọc. Bài viết phân tích các tính chất đặc trưng của các phép lọc tuyến tính phổ biến: Phép lấy sai phân, phép trung bình trượt, phép lọc thông cao, thông thấp, thông dải, Hodrick-Prescott và Baxter-King thông qua phân tích hàm truyền và hàm lợi ích. Với mục đích tách thành phần xu hướng của một chuỗi thời gian, chỉ các phép lọc thông cao, Hodrick-Prescott và Baxter-King có thể được sử dụng. Bên cạnh đó, ba loại phép lọc này không làm thay đổi pha và biên độ dao động của thành phần chu kỳ so với chuỗi dữ liệu gốc. Từ một nghiên cứu thực nghiệm các phép lọc thông cao, Hodrick-Prescott, Baxter-King trên chuỗi chỉ số giá chứng khoán VN-Index tần số tuần, nhóm tác giả đã tìm được các tham số hợp lý cho các phép lọc thông cao, Hodrick-Prescott và Baxter-King. Abstract Filters are the popular tools for analyzing a time series into trend and cyclical components. The common property of the filters is retaining some components of the original time series as well as affecting the amplitude and the phase of the series received after filtering. The article analyzes the characteristic properties of popular linear filters: Differences, moving averages, high-pass, low-pass, band-pass, Hodrick-Prescott and Baxter-King filters through analysis of the

* Tác giả liên hệ.

Email: mybtt@buh.edu.vn (Bùi Thị Thiên Mỹ), yennt@buh.edu.vn (Nguyễn Thị Yến).

Trích dẫn bài viết: Bùi Thị Thiên Mỹ, & Nguyễn Thị Yến. (2022). Phép lọc tuyến tính và vấn đề khử xu hướng của chuỗi thời gian: Nghiên cứu thực nghiệm tại Việt Nam. *Tạp chí Nghiên cứu Kinh tế và Kinh doanh Châu Á*, 33(5), 66–83.

transfer functions and the gain functions. For the purpose of extracting the trend component, only high-pass (HF), Hodrick-Prescott (HPF) and Baxter-King (BKF) filters can be used. Besides, these three types of filters do not change the phase and the amplitude of the cyclical component compared to the original data series. With the experimental study of HF, HPF, BKF filters on the weekly frequency VN-Index, the authors find out the reasonable parameters for HF, HPF, and BKF.

1. Giới thiệu

Trong nghiên cứu chuỗi thời gian, hai kiểu phân tích thường được sử dụng là phân tích trên miền thời gian (Time Series Analysis), và phân tích trên miền tần số (Frequency Analysis) (còn được gọi là phân tích phổ (Spectral Analysis)). Trong đó, phân tích phổ đòi hỏi nhiều kiến thức toán học chuyên sâu và kỹ năng tính toán phức tạp. Tuy vậy, phân tích phổ giúp nghiên cứu tăng sâu các thành phần dao động tương ứng với các tần số khác nhau của chuỗi thời gian. Trong nghiên cứu kinh tế vĩ mô, một số giả thuyết về mối quan hệ giữa các yếu tố được xây dựng dựa trên phân tích phổ có thể liệt kê như: Giả thuyết về độ dốc của đường cong Philips thể hiện mối quan hệ giữa lạm phát và thất nghiệp trong ngắn hạn (ngược chiều) và dài hạn (cùng chiều) (Iacobucci, 2005); giả thuyết về ảnh hưởng của cung tiền đến lạm phát trong ngắn hạn và dài hạn (Kamalian và cộng sự, 2020; Tastan & Sahin, 2020). Một kết quả điển hình của phân tích phổ là lý thuyết chu kỳ kinh doanh của nền kinh tế – quan tâm đến các thành phần có chu kỳ từ 1,5 đến 8 năm, trong khi đó, lý thuyết tăng trưởng kinh tế – quan tâm đến các thành phần dao động có chu kỳ rất dài (Hodrick & Prescott, 1997; Hornsterin, 1998).

Trong phân tích phổ, phép lọc được sử dụng nhằm tách chuỗi thời gian thành các thành phần dao động có chu kỳ ngắn (thành phần chu kỳ (Cyclical Component)) và chu kỳ dài (thành phần xu hướng (Trend Component)). Các phép lọc tuyến tính phổ biến là các phép lấy sai phân, các phép trung bình trượt, phép lọc thông cao (High-Pass Filter – HF), phép lọc Hodrick-Prescott (HPF) và phép lọc Baxter-King (BKF). Sau phép lọc, chuỗi dữ liệu đầu ra bị loại bỏ hoặc khuếch đại một số thành phần dao động tương ứng với tần số nào đó, đồng thời các mối quan hệ về thời gian của các thành phần trong chuỗi dữ liệu ban đầu có thể bị thay đổi. Mặc dù vậy, các nghiên cứu thường sử dụng phép lọc một cách chủ quan và không đánh giá đầy đủ mối quan hệ giữa chuỗi trước và sau phép lọc (Hamilton, 2018). Điều này có thể dẫn đến sự mạo nhận trong các kết quả suy diễn. Tại Việt Nam, phép sai phân thường xuyên được áp dụng với mục đích dừng hóa chuỗi thời gian. Tuy nhiên, phép sai phân cho ra chuỗi có phương sai lớn hơn rất nhiều lần so với chuỗi gốc. Một lựa chọn khác là HPF, được áp dụng trên các chuỗi dữ liệu tần số tháng, quý và năm với tham số theo đề xuất từ những nghiên cứu gốc tại Hoa Kỳ (Baxter & King, 1999; Hodrick & Prescott, 1997). Điều này gây khó khăn khi nghiên cứu các chuỗi thời gian có tần số cao hơn trong bối cảnh Việt Nam. Vì những lý do trên, bài báo thảo luận các tính chất của phép lọc dưới góc độ lý thuyết phân tích phổ, từ đó nhấn mạnh phạm vi áp dụng để nâng cao độ tin cậy của các kết quả phân tích chuỗi thời gian dựa trên phép lọc. Bên cạnh đó, một tình huống thực nghiệm trên chuỗi chỉ số giá chứng khoán VN-Index được tiến hành nhằm minh họa quy trình ước lượng các tham số cho HF, HPF và BKF đối với dữ liệu tần số tuần.

Phần tiếp theo của bài báo được bố cục như sau: Phần 2 nhắc lại một số khái niệm và kiến thức liên quan đến phép lọc. Phần 3 là những phân tích các phép lọc từ góc nhìn lý thuyết. Phần 4 là kết quả thực nghiệm các phép lọc trên chuỗi VN-Index. Phần 5 là những kết luận.

2. Các kiến thức liên quan

Khái niệm “dừng” đối với các chuỗi thời gian trong bài viết được hiểu là dừng hiệp phương sai (Covariance Stationary). Các định nghĩa và định lý bên dưới được tham khảo trong các tài liệu kinh tế lượng (Hamilton, 2020; Neusser, 2016).

2.1. Phép lọc tuyến tính

- **Định nghĩa 2.1:** Xét không gian Ω gồm các chuỗi thời gian.

Ánh xạ $L: \Omega \rightarrow \Omega$ sao cho $LX_t = X_{t-1}, \forall X_t \in \Omega$ được gọi là toán tử trễ.

Toán tử $\psi(L) = \sum_{j=-\infty}^{+\infty} \psi_j L^j$ (với $\sum_{j=-\infty}^{+\infty} |\psi_j| < +\infty$) được gọi là phép lọc tuyến tính trên không gian Ω . Nói cách khác, phép lọc tuyến tính $\psi(L)$ là ánh xạ $\psi(L): \Omega \rightarrow \Omega$, sao cho: $\psi(L)X_t = \sum_{j=-\infty}^{+\infty} \psi_j L^j X_t = \sum_{j=-\infty}^{+\infty} \psi_j X_{t-j}, \forall X_t \in \Omega$.

Chuỗi thời gian $\{Y_t\}$ với $Y_t = \sum_{j=-\infty}^{+\infty} \psi_j X_{t-j}$ được gọi là đầu ra (ảnh) của chuỗi thời gian $\{X_t\}$ qua phép lọc tuyến tính $\psi(L)$.

- **Định nghĩa 2.2:** Cho $\{X_t\}$ là một quá trình dừng với trung bình bằng 0 và $\gamma(h)$ là hàm tự hiệp phương sai: $\gamma(h) = \text{cov}(X_t, X_{t+h})$ thỏa $\sum_{h=-\infty}^{+\infty} \gamma(h) < +\infty$. Khi đó, hàm số

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} \gamma(h) e^{-ih\lambda}, -\infty < \lambda < +\infty$$

được gọi là “hàm mật độ phổ” của quá trình $\{X_t\}$.

- **Định lý 2.1:** Cho $\{X_t\}$ là một quá trình dừng. Khi đó,

$$X_t = \mu + \int_0^\pi (\alpha_1(\lambda) \cos \lambda t + \alpha_2(\lambda) \sin \lambda t) d\lambda$$

Ở đây, $\alpha_i(\cdot)$ với $i = 1, 2$ là các biến ngẫu nhiên có trung bình bằng 0, và bất kỳ các tần số $0 < \omega_1 < \omega_2 < \omega_3 < \omega_4 < \pi$ thì:

$$\int_{\omega_1}^{\omega_2} \alpha_i(\lambda) d\lambda \text{ không tương quan với } \int_{\omega_3}^{\omega_4} \alpha_i(\lambda) d\lambda, i = 1, 2$$

Bất kỳ $0 < \omega_1 < \omega_2 < \pi$ và $0 < \omega_3 < \omega_4 < \pi$, $\int_{\omega_1}^{\omega_2} \alpha_1(\lambda) d\lambda$ không tương quan với $\int_{\omega_3}^{\omega_4} \alpha_2(\lambda) d\lambda$.

Như vậy, theo Định lý 2.1, quá trình ngẫu nhiên dừng $\{X_t\}$ là tổng của các chuỗi điều hòa có tần số thuộc đoạn $[0, \pi]$. Khi muốn chỉ thành phần dao động tương ứng với tần số cụ thể λ_0 của $\{X_t\}$, ta viết $X(\lambda_0)$.

- **Định lý 2.2:** Cho $\{X_t\}$ là một quá trình dừng với trung bình bằng 0. Khi đó, nếu quá trình $\{Y_t\}$ là đầu ra của $\{X_t\}$ qua phép lọc tuyến tính $\psi(L)$ thì $\{Y_t\}$ là một quá trình dừng với trung bình bằng 0. Hơn nữa, hàm mật độ phổ của các chuỗi $\{X_t\}$ và $\{Y_t\}$ có mối quan hệ:

$$f_Y(\lambda) = \psi(e^{-i\lambda})\psi(e^{i\lambda})f_X(\lambda)$$

Trong đó, $\psi(e^{-i\lambda}) = \sum_{j=-\infty}^{+\infty} \psi_j e^{-ij\lambda}$.

- **Định nghĩa 2.3:** Với mô tả như Định lý 2.2, hàm $TF(\lambda) = \psi(e^{-i\lambda}) = \sum_{j=-\infty}^{+\infty} \psi_j e^{-ij\lambda}$ được gọi là “hàm truyền” (Transfer Function) hoặc “hàm đáp ứng tần số” (Frequency-Response Function) của phép lọc $\psi(L)$. Hàm $G(\lambda) = \psi(e^{-i\lambda})\psi(e^{i\lambda})$ được gọi là “hàm lợi ích” (Gain Function) của phép lọc.

2.2. Các tính chất của phép lọc tuyến tính

Giả sử, $\{Y_t\}$ là ảnh của $\{X_t\}$ qua phép lọc $\psi(L)$: $Y_t = \psi(L)X_t$.

Từ Định lý 2.2, trong miền tần số, ta có: $Y(\lambda) = TF(\lambda)X(\lambda)$, với $\lambda \in [0, \pi]$.

2.2.1. Vai trò của hàm truyền trong phép lọc

Vì $\arg(Y(\lambda)) = \arg(TF(\lambda)) + \arg(X(\lambda))$ nên sau phép lọc, chuỗi $\{Y_t\}$ bị thay đổi argument¹ so với chuỗi $\{X_t\}$. Thuật ngữ mô tả hiện tượng này là “phép lọc làm dịch chuyển pha của chuỗi ban đầu”. Nghĩa là, một sự kiện xảy ra ở thời điểm t được biểu diễn bởi chuỗi thời gian $\{X_t\}$ qua phép lọc sẽ không xảy ra ở thời điểm t nữa.

Chuỗi $\{Y_t\}$ không đổi pha so với $\{X_t\}$ khi và chỉ khi $TF(\lambda) = TF(-\lambda), \forall \lambda \in [0, \pi]$.

Bên cạnh đó, có thể xảy ra các trường hợp sau đây: Xét $\lambda \in [0, \pi]$

- $TF(\lambda) = 0$ với $\lambda \geq \lambda_0$. Nghĩa là, sau phép lọc, các thành phần có tần số cao trong $\{X_t\}$ đã bị loại bỏ, $\{Y_t\}$ chỉ chứa các thành phần với tần số thấp hơn λ_0 . Khi đó, phép lọc được gọi là phép lọc “thông thấp” (Low-Pass Filter).

- $TF(\lambda) = 0$ với $\lambda \leq \lambda_0$. Khi đó, chuỗi $\{Y_t\}$ chỉ chứa các thành phần có tần số lớn hơn λ_0 . Khi đó, phép lọc được gọi là phép lọc “thông cao” (High-Pass Filter).

- $TF(\lambda) \neq 0$ với $0 \leq \lambda_1 \leq \lambda \leq \lambda_2$. Khi đó, phép lọc được gọi là phép lọc “thông dải” (Band-Pass Filter), $\{Y_t\}$ chỉ chứa các thành phần có tần số nằm trong khoảng λ_1 đến λ_2 .

Các giá trị $\lambda_0, \lambda_1, \lambda_2$ tương ứng với các trường hợp trên được gọi là tần số cắt (tần số ngưỡng).

2.2.2. Vai trò của hàm lợi ích trong phép lọc

Vì $|Y(\lambda)| = |TF(\lambda)| \cdot |X(\lambda)|$, và $G(\lambda) = |TF(\lambda)|^2$ nên hàm lợi ích chỉ sự thay đổi biên độ dao động của chuỗi $\{Y_t\}$ so với chuỗi $\{X_t\}$. Nếu $G(\lambda_0) > 1$, phép lọc sẽ khuếch đại dao động của thành phần tương ứng với tần số λ_0 trong chuỗi $\{X_t\}$. Ngược lại, nếu $G(\lambda_0) < 1$, phép lọc nén thành phần dao động tương ứng với tần số λ_0 . Trong ứng dụng phép lọc, người ta kỳ vọng biên độ dao động của chuỗi đầu ra không thay đổi so với chuỗi gốc, nghĩa là $G(\lambda) = 1$ với λ thuộc khoảng tần số mà phép lọc giữ lại.

Phép lọc được gọi là “lý tưởng” nếu $\{Y_t\}$ có pha và biên độ không thay đổi so với $\{X_t\}$.

¹ Một số phức $z = a + ib$ ($a, b \in \mathbb{R}$) có dạng lượng giác là $z = r(\cos \varphi + i \sin \varphi)$ với $r = \sqrt{a^2 + b^2}$, $\tan \varphi = \frac{b}{a}$ thì r được gọi là modun, φ được gọi là argument của số phức z .

3. Góc nhìn từ lý thuyết các phép lọc

Phần này phân tích tính chất một số phép lọc phổ biến trong nghiên cứu kinh tế thông qua phân tích hàm truyền và hàm lợi ích.

3.1. Phép lấy sai phân

3.1.1. Sai phân bậc nhất và sai phân bậc hai

- Chuỗi $\{Y_t\}$ được gọi là “chuỗi sai phân bậc nhất” liên kết với chuỗi $\{X_t\}$ nếu:

$$Y_t = X_t - X_{t-1} = (1 - L)X_t$$

Hàm truyền và hàm lợi ích của phép lấy sai phân bậc nhất là:

$$TF_1(\lambda) = 1 - e^{-i\lambda}, G_1(\lambda) = 2(1 - \cos\lambda)$$

- Chuỗi $\{Y_t\}$ được gọi là “chuỗi sai phân bậc hai” liên kết với chuỗi $\{X_t\}$ nếu:

$$Y_t = (1 - L)^2 X_t = [X_t - X_{t-1}] - [X_{t-1} - X_{t-2}]$$

Hàm truyền và hàm lợi ích của phép lấy sai phân bậc hai là:

$$TF_2(\lambda) = (1 - e^{-i\lambda})^2, G_2(\lambda) = 4(1 - \cos\lambda)^2$$

Bảng 1 thể hiện giá trị của hàm lợi ích $G_1(\lambda)$ và $G_2(\lambda)$ tương ứng với các tần số khác nhau. Các phép lấy sai phân bậc nhất và bậc hai loại bỏ các thành phần dao động có tần số rất thấp (≈ 0) trong chuỗi $\{X_t\}$ (tương đương loại bỏ các thành phần có chu kỳ dài trong chuỗi ban đầu). Tuy vậy, các phép lấy sai phân này có thể khuếch đại một số thành phần của $\{X_t\}$. Các thành phần có tần số càng cao ($\approx \pi$) thì giá trị hàm lợi ích càng lớn, nghĩa là biên độ dao động của chúng được phóng đại nhiều lần. Điều đó có nghĩa là phép lấy sai phân đã phóng đại các thành phần bất quy tắc (thành phần nhiễu, thành phần có tần số cao) của $\{X_t\}$ và làm tăng phương sai của chuỗi đầu ra $\{Y_t\}$ so với phương sai của chuỗi gốc $\{X_t\}$. Như vậy, về mặt thực hành, nếu chuỗi $\{X_t\}$ có tần số cao (ví dụ: đơn vị tuần, ngày, giờ) thì không nên sử dụng các phép sai phân để khử thành phần xu hướng của $\{X_t\}$. Tuy nhiên, nếu $\{X_t\}$ là chuỗi có tần số thấp (ví dụ: đơn vị là tháng, quý, năm) thì các phép sai phân có thể cân nhắc áp dụng.

Bảng 1.

Giá trị hàm lợi ích của các phép sai phân bậc nhất và bậc hai

λ	0	$\frac{\pi}{3}$	$\frac{\pi}{2}$	π
$G_1(\lambda)$	0	1	2	4
$G_2(\lambda)$	0	1	4	16

3.1.2. Phép lấy sai phân mùa

Phép khử tính mùa thường được thực hiện bằng phép lấy sai phân theo tần số quan sát trong năm của dữ liệu, gọi là phép lấy “sai phân mùa”. Phép lấy sai phân mùa tần số quan sát k lần/năm là:

$$Y_t = (1 - L^k)X_t = S_k(L)X_t, \text{ với } k \in \mathbb{N}^*.$$

Hàm truyền và hàm lợi ích của phép lọc sai phân mùa $S_k(L) = 1 - L^k$ là:

$$TF^{S_k}(\lambda) = 1 - e^{-ik\lambda}, G^{S_k}(\lambda) = 2(1 - \cos k\lambda)$$

Phép lọc $S_k(L)$ loại bỏ các thành phần có chu kỳ quá dài và các thành phần có chu kỳ $k, \frac{k}{2}, \frac{k}{3}, \dots \in [0, \pi]$ trong chuỗi gốc.

Có thể thấy các phép lấy sai phân (bậc nhất, bậc hai, mùa) có tính chất của một phép lọc thông cao: Loại bỏ được thành phần có chu kỳ dài. Tuy nhiên, vì các hàm truyền không là hàm số chẵn nên các phép lấy sai phân làm thay đổi pha của chuỗi đầu ra so với chuỗi gốc. Mặt khác, các phép lấy sai phân đã phóng đại các thành phần bất quy tắc và phóng đại phương sai của chuỗi ban đầu.

3.2. Phép trung bình trượt

Phép lọc trung bình trượt được mô tả như sau:

$$Y_t = \sum_{j=H}^K \psi_j X_{t-j} = \psi(L)X_t, \text{ với } \psi(L) = \sum_{j=H}^K \psi_j L^j.$$

Phép trung bình trượt được gọi là “đối xứng” nếu $H = -K$ và $\psi_j = \psi_{-j}, \forall j$.

3.2.1. *Phép trung bình trượt trung tâm có các trọng số bằng nhau (Central Moving Average – CMA)*

Khi đó: $H = -K$ và $\psi_j = \frac{1}{2K+1} (j = -K, \dots, K)$

Hàm truyền và hàm lợi ích tương ứng là:

$$TF(\lambda) = \frac{1}{2K+1} [\sum_{j=1}^K (2\cos j\lambda) + 1], G(\lambda) = TF(\lambda)^2, \text{ với } \lambda \in [0, \pi].$$

Trường hợp này, phép lọc CMA loại bỏ các thành phần có tần số dao động là bội khác 0 của $\frac{2\pi}{2K+1}$. Đối với thành phần có tần số rất bé, CMA không đổi biên độ dao động. Như vậy, CMA không loại bỏ xu hướng của các thành phần có chu kỳ dài trong chuỗi $\{X_t\}$. Mặt khác, CMA không làm thay đổi pha của chuỗi ban đầu. Các phép lọc CMA là các phép lọc thông thấp.

3.2.2. *Phép trung bình trượt đối xứng và tổng các trọng số bằng 0*

Vì $\sum_{j=-K}^K \psi_j = 0$ nên nếu λ bằng 0 thì hàm truyền bằng 0. Phép lọc này khử được thành phần có chu kỳ rất dài và giữ không đổi pha chuỗi đầu ra.

3.2.3. *Xấp xỉ phép lọc thông thấp lý tưởng bằng phép trung bình trượt đối xứng*

Nhắc lại, “phép lọc thông thấp lý tưởng” (LF_{ideal}) là phép lọc thông thấp và không làm thay đổi biên độ dao động của chuỗi đầu ra. Hàm truyền của phép lọc thông thấp lý tưởng là:

$$TF_{\text{thấp}}(\lambda) = \begin{cases} 1 & \text{nếu } |\lambda| \leq \lambda_0 \\ 0 & \text{nếu } |\lambda| > \lambda_0 \end{cases} \text{ với } \lambda \in [-\pi, \pi] \quad (1)$$

với λ_0 là tần số cắt của phép lọc thông thấp lý tưởng. Trong miền thời gian, LF_{ideal} được viết dưới dạng trung bình trượt đối xứng vô hạn: $LF_{\text{ideal}}(L) = \sum_{j=-\infty}^{+\infty} b_j L^j$.

Thực hiện biến đổi Fourier ngược đối với hàm truyền cho bởi phương trình (1), ta có:

$$b_j = \begin{cases} \lambda_0/\pi & \text{nếu } j = 0 \\ \sin(j\lambda_0)/j\pi & \text{nếu } j \neq 0 \end{cases} \quad (2)$$

Theo công thức (2), khi $j \rightarrow +\infty$ thì $b_j \rightarrow 0$. Vì thế, có thể xây dựng một phép lọc thông thấp tần số cắt là λ_0 bằng phép trung bình trượt hữu hạn $LF(L) = \sum_{j=-K}^K a_j L^j$, với K đủ lớn và $a_j = b_j, |j| \leq K$.

Phép lọc thông thấp $LF(L)$ là một xấp xỉ của phép lọc thông thấp lý tưởng LF_{ideal} . Khi K tăng, kết quả xấp xỉ tốt hơn. Tuy nhiên, chuỗi dữ liệu đầu ra mất nhiều quan sát hơn (số quan sát bị mất là $2K$). Vì thế, giá trị K được lựa chọn phụ thuộc vào độ dài chuỗi dữ liệu và yêu cầu độ chính xác của phép xấp xỉ.

3.2.4. Xây dựng các phép lọc thông cao và thông dải từ phép trung bình trượt đối xứng

Khi xử lý hướng một chuỗi thời gian bằng phép lọc, cần thiết kế phép lọc để có thể loại bỏ các thành phần dao động với tần số rất bé của chuỗi ban đầu. Phép lọc thông cao và phép lọc thông dải là những lựa chọn đáp ứng được yêu cầu đó. Hàm truyền của phép lọc thông cao lý tưởng và phép lọc thông dải lý tưởng là:

$$TF_{cao}(\lambda) = \begin{cases} 1 & \text{nếu } |\lambda| \geq \lambda_0 \\ 0 & \text{nếu } |\lambda| < \lambda_0 \end{cases} \quad (3)$$

$$TF_{dải}(\lambda) = \begin{cases} 1 & \text{nếu } \lambda_1 \leq |\lambda| \leq \lambda_2 \\ 0 & \text{nếu } |\lambda| < \lambda_1 \text{ hoặc } |\lambda| > \lambda_2 \end{cases} \quad (4)$$

Khi đó, với cùng tần số cắt λ_0 , ta có mối quan hệ: $TF_{cao} = 1 - TF_{thấp}$.

Ký hiệu $TF_{thấp}^{(1)}$ và $TF_{thấp}^{(2)}$ là hàm truyền của các phép lọc thông thấp lý tưởng với các tần số cắt là λ_1 và λ_2 . Khi đó, hàm truyền của phép lọc thông dải lý tưởng được xác định như sau:

$$TF_{dải} = TF_{thấp}^{(2)} - TF_{thấp}^{(1)}$$

Từ các mối quan hệ này, ta có thể xây dựng được các xấp xỉ phép lọc thông cao và phép lọc thông dải lý tưởng.

3.3. Phép lọc Hodrick-Prescott (HPF)

Phép lọc Hodrick-Prescott – HPF (Hodrick & Prescott, 1997) được xây dựng dựa trên giả định chuỗi $\{X_t\}$ được phân rã thành hai yếu tố: Thành phần xu hướng $\{G_t\}$ và thành phần chu kỳ $\{C_t\}$ như sau: $X_t = G_t + C_t$

3.3.1. Bài toán tìm phép lọc Hodrick-Prescott

Hodrick và Prescott (1997) đề xuất một phương pháp tìm thành phần $\{G_t\}$ như sau:

$$G_t^{HP} = \min_{G_t} \{ \sum_{t=1}^T (X_t - G_t)^2 + \lambda \sum_{t=2}^{T-1} [(G_{t+1} - G_t) - (G_t - G_{t-1})]^2 \} \quad (5)$$

Hàm mục tiêu trong bài toán tối ưu (5) có hai thành phần. Thành phần $\sum_{t=1}^T (X_t - G_t)^2$ đo sự tương thích giữa chuỗi dữ liệu ban đầu $\{X_t\}$ và thành phần xu hướng $\{G_t\}$. Thành phần $\sum_{t=2}^{T-1} [(G_{t+1} - G_t) - (G_t - G_{t-1})]^2$ đo độ trơn của chuỗi $\{G_t\}$. Bài toán tối ưu (5) đứng giữa sự đánh đổi: Sự tương thích giữa $\{G_t\}$ và $\{X_t\}$ với độ trơn của chuỗi $\{G_t\}$. Sự đánh đổi này được điều chỉnh bằng tham số làm trơn (Smoothness Parameter) $\lambda > 0$.

Lời giải của bài toán (5) được xác định như sau:

$$G_t^{HP} = \frac{1}{\lambda L^{-2}(1-L)^4 + 1} X_t; \quad C_t^{HP} = \frac{\lambda L^{-2}(1-L)^4}{\lambda L^{-2}(1-L)^4 + 1} X_t \quad (6)$$

Trong miền thời gian, G_t^{HP} là kết quả của một phép lọc trung bình trượt vô hạn (Baxter & King, 1999). Vì thế, HPF là một phép lọc tuyến tính.

Khi đó, công thức của HPF là: $HP(L) = \frac{\lambda L^{-2}(1-L)^4}{\lambda L^{-2}(1-L)^4 + 1}$

Hàm truyền của HPF là: $TF^{HP}(\omega) = \frac{4\lambda(1-\cos\omega)^2}{4\lambda(1-\cos\omega)^2 + 1}$, với ω chỉ tần số.

3.3.2. Tính chất của phép lọc Hodrick-Prescott

- $TF^{HP}(\omega)$ là hàm số chẵn trên $[-\pi; \pi]$ nên HPF không làm thay đổi pha của chuỗi gốc.

- Xét $\omega \in [0; \pi]$. $TF^{HP}(\omega) = 0$ khi và chỉ khi $\omega = 0$, nên HPF loại các thành phần có tần số rất bé. Đây là một đặc điểm của phép lọc thông cao.

- $|TF^{HP}(\omega)| < 1, \forall \omega \in [-\pi; \pi]$. Khi λ khá lớn, $|TF^{HP}(\pi)| \approx 1$ nên HPF gần như giữ nguyên biên độ dao động của các thành phần có tần số cao. Đặc điểm này khiến HPF gần giống với một phép lọc thông cao lý tưởng.

3.3.3. Những lưu ý khi sử dụng phép lọc Hodrick-Prescott

(1) Thành phần xu hướng được xác định bởi công thức (6) phụ thuộc vào giá trị tham số làm trơn λ . Nghiên cứu thực nghiệm trên chuỗi GDP của Hoa Kỳ, Hodrick và Prescott (1997) đề xuất $\lambda = 1.600$ đối với dữ liệu quý. Khi đó, HPF loại các thành phần có tần số nhỏ hơn $\frac{\pi}{16}$ (tương đương với chu kỳ lớn hơn 32 quý). Hình 1 thể hiện đồ thị hàm truyền của HPF với dữ liệu lấy theo quý và năm, tần số cắt tương đương là 8 năm.

(2) Mặc dù có những tranh cãi về giá trị của tham số làm trơn λ , một đề xuất của Ravn và Uhlig (2002) thường được áp dụng trong thực hành là:

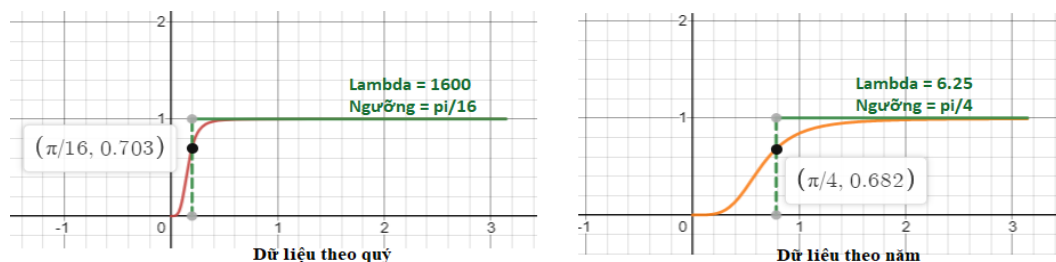
$$\lambda_s = \left(\frac{s}{4}\right)^m \lambda_4 \quad (7)$$

Trong đó, λ_4, λ_s lần lượt là tham số làm phẳng cho dữ liệu quý và dữ liệu lấy tần suất s lần/năm; m có thể nhận giá trị 3,8 hoặc 4 (Ravn & Uhlig, 2002). Nếu chọn $\lambda_4 = 1.600$ và $m = 4$ thì giá trị của tham số làm phẳng với các dữ liệu lấy theo tần suất khác nhau được xác định như sau:

$$\lambda = \begin{cases} 6,25 & \text{nếu } s = 1 \\ 1600 & \text{nếu } s = 4 \\ 129600 & \text{nếu } s = 12 \end{cases}$$

(3) HPF tạo ra thành phần chu kỳ bị sai lệch tương ứng với các quan sát đầu và cuối chuỗi dữ liệu. Khi đó, cách đơn giản nhất để khắc phục là áp dụng HPF cho toàn bộ mẫu, nhưng sau đó, loại bỏ một số quan sát ở đầu và cuối mẫu (Baxter & King, 1999).

(4) HPF cho kết quả tốt với các chuỗi dữ liệu có xu hướng tương đối ổn định. Tuy vậy, HPF có thể sinh ra các mối quan hệ động giả mạo giữa các chuỗi dữ liệu khi các chuỗi ban đầu không tách bạch rõ thành phần xu hướng và chu kỳ (Dadashova, 2012; Hamilton, 2018).



Hình 1. Đồ thị hàm truyền của HPF với ngưỡng cắt 8 năm

3.4. Phép lọc Baxter-King (BKF)

3.4.1. Ý nghĩa và công thức của phép lọc Baxter-King

Dựa trên quan sát về chu kỳ kinh doanh thực, Baxter và King (1999) cho rằng khi nghiên cứu một chuỗi thời gian kinh tế, cần tách riêng thành phần chu kỳ ra khỏi các thành phần như: Thành phần xu hướng – thay đổi rất chậm, thành phần mùa vụ, và thành phần bất quy tắc – thay đổi nhanh. Nói cách khác, chu kỳ kinh doanh thực là thành phần gồm các dao động có chu kỳ không vượt quá p_2 và không bé hơn p_1 ($p_1 < p_2$).

Dựa trên ý tưởng đó, Baxter và King xây dựng một phép lọc thông dải xấp xỉ phép lọc thông dải lý tưởng bằng phép trung bình trượt đối xứng và tổng các trọng số bằng 0, ký hiệu là $BK_K(p_1, p_2)$.

$$BK_K(p_1, p_2) = \sum_{j=-K}^K a_j L^j$$

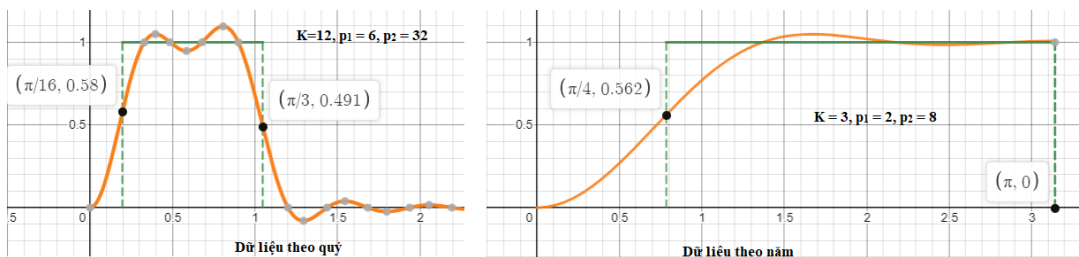
Các trọng số a_j được tính theo công thức sau:

$$\begin{aligned} a_j &= b_j + \theta \\ b_j &= \begin{cases} 2 \left(\frac{1}{p_1} - \frac{1}{p_2} \right) & \text{nếu } j = 0 \\ [\sin(2j\pi/p_1) - \sin(2j\pi/p_2)] \frac{1}{j\pi} & \text{nếu } j = \pm 1, \pm 2, \dots \end{cases} \\ \theta &= \frac{-\sum_{j=-K}^K b_j}{2K + 1} \end{aligned} \quad (8)$$

3.4.2. Tính chất và những chú ý khi sử dụng phép lọc Baxter-King

BKF không làm thay đổi pha và biên độ dao động của chuỗi đầu ra so với chuỗi ban đầu.

Dựa trên dữ liệu kinh tế của Hoa Kỳ, Baxter và King (1999) đề nghị K nên được chọn tương đương với phép trung bình trượt tính trên ba năm quá khứ và ba năm tương lai. Cụ thể, đối với dữ liệu quý, $K = 12$, $p_1 = 6$, $p_2 = 32$, phép lọc thông dải $BK_{12}(6,32)$ tách riêng thành phần dao động có chu kỳ nằm trong khoảng 6 đến 32 quý. Đối với dữ liệu năm, $K = 3$; $p_1 = 2$; $p_2 = 8$, phép lọc thông dải $BK_3(2,8)$ tách riêng thành phần dao động có chu kỳ trong khoảng từ 2 đến 8 năm. Tuy nhiên, đối với các nghiên cứu thực nghiệm, giá trị K còn tùy thuộc vào độ dài của chuỗi dữ liệu và yêu cầu chính xác của phép xấp xỉ với phép thông dải lý tưởng. Hình 2 thể hiện đồ thị hàm truyền của BKF cho dữ liệu quý và năm theo Baxter-King.



Hình 2. Đồ thị hàm truyền của BKF cho dữ liệu quý và năm

4. Góc nhìn từ thực nghiệm một số phép lọc

Trong những nghiên cứu trước đây (Baxter & King, 1999; Hodrick & Prescott, 1997; Ravn & Uhlig, 2002), các phép lọc thường áp dụng cho dữ liệu kinh tế vĩ mô Hoa Kỳ theo năm, quý và tháng. Vì thế, tham số của các phép lọc HPF và BKF được đề xuất dựa vào chu kỳ kinh doanh của Hoa Kỳ và tần số của dữ liệu mẫu. Tại Việt Nam, sử dụng phép lọc trên chuỗi dữ liệu có tần số cao hơn chưa được quan tâm. Do đó, nhóm tác giả khảo sát một số phép lọc trên chuỗi chỉ số giá chứng khoán VN-Index tần số tuần. Từ kết quả phân tích, nhóm tác giả mong muốn bổ sung thêm sự lựa chọn các tham số cho HF, HPF và BKF trên các chuỗi dữ liệu tuần tại Việt Nam.

4.1. Dữ liệu thực nghiệm

4.1.1. Thống kê mô tả

Chuỗi thời gian được sử dụng là chuỗi logarit tự nhiên của chỉ số giá chứng khoán VN-Index (ký hiệu LVNI). Trong đó, chuỗi VN-Index được thu thập theo tuần, từ tuần lễ đầu tiên đến ngày 12/05/2021, với tổng số quan sát là 1.062 tuần. Nhằm tránh hiệu ứng ngày đầu tuần và ngày cuối tuần, mẫu ghi nhận chỉ số đóng cửa vào thứ Tư. Nếu thứ Tư không là ngày giao dịch, chỉ số đóng cửa thứ Năm được chọn thay thế. Nếu cả ngày thứ Tư và thứ Năm không giao dịch, trung bình chỉ số các ngày còn lại trong tuần được chọn. Kết quả kiểm định nghiệm đơn vị cho thấy chuỗi LVNI dừng với mức ý nghĩa 5%. Bảng 2 trình bày các đặc trưng thống kê mô tả của chuỗi LVNI và kết quả kiểm định nghiệm đơn vị.

Bảng 2.

Thống kê mô tả và kết quả kiểm định nghiệm đơn vị chuỗi LVNI

Thống kê mô tả	Số quan sát	1.062
	Trung bình	6,15
	Độ lệch chuẩn	0,58
	Giá trị nhỏ nhất	4,64
	Giá trị lớn nhất	7,15
Kiểm định nghiệm đơn vị Dickey–Fuller	Thống kê kiểm định	–1,99
	Giá trị tới hạn 1%	–2,33
	Giá trị tới hạn 5%	–1,65
	Giá trị tới hạn 10%	–1,28

4.1.2. Chu kỳ của chuỗi VN-Index

Theo dõi biến động của chuỗi VN-Index từ năm 2000 đến tháng 5 năm 2021, nhận thấy VN-Index đã trải qua sáu chu kỳ tăng giảm. Khoảng cách giữa hai đáy của các chu kỳ liên tiếp nhau từ 30 tháng đến 42 tháng. Từ đó, có thể nhận định rằng, trung bình độ dài một chu kỳ của chuỗi VN-Index là khoảng ba năm, tính theo tuần là 156 tuần (tương đương tần số là $\lambda_0 = \frac{2\pi}{156}$). Bảng 3 tóm tắt thông tin các chu kỳ của chuỗi VN-Index.

Bảng 3.

Các chu kỳ của chuỗi VN-Index

Chu kỳ	Đáy bắt đầu	Đáy kết thúc	Thời gian giữa hai đáy
1	24/10/2003	02/8/2006	33 tháng
2	02/8/2006	24/02/2009	30 tháng
3	24/02/2009	06/01/2012	34 tháng
4	06/01/2012	17/12/2014	35 tháng
5	17/12/2014	11/7/2018	42 tháng
6	11/7/2018	10/02/2021	30 tháng

4.2. Các kết quả thực nghiệm

Từ những phân tích ở mục 3, một số tính chất quan trọng của các phép lọc được tóm tắt ở Bảng 4. Với mục đích tách thành phần xu hướng của chuỗi LVNI, đồng thời không làm thay đổi pha, biên độ dao động của thành phần chu kỳ, các phép lọc HF, HPF và BKF nên được chọn. Các phép sai phân, trung bình trượt và thông thấp không thỏa mãn các yêu cầu về pha, biên độ dao động và khả năng tách thành phần có chu kỳ dài, vì thế, không được áp dụng trong phần nghiên cứu thực nghiệm này.

Các phép lọc HF, BKF phụ thuộc vào giá trị K toán tử trễ và toán tử tiến trong phép trung bình trượt. Vì thế, nhóm tác giả thử nghiệm phép lọc đối với các giá trị K khác nhau. Sau đó, căn cứ vào sự ổn định của độ dao động (Volatility) (đo bằng độ lệch chuẩn) và sự ổn định của độ bền (Persistence) (đo bằng hệ số tự tương quan bậc một), giá trị K hợp lý sẽ được chọn. Phương pháp xác định K như trên đã được Baxter và King (1999) sử dụng.

Đối với HPF, kết quả phép lọc phụ thuộc vào tham số làm trơn λ . Nhóm tác giả đề xuất một cách ước lượng λ . Việc chọn giá trị λ tùy thuộc vào yêu cầu xấp xỉ phép lọc thông cao lý tưởng.

Bảng 4.

Các tính chất của một số phép lọc

Thứ tự	Phép lọc	Thay đổi pha	Khuếch đại biên độ	Khử thành phần chu kỳ dài
1	Sai phân bậc một	Có	Có	Có
2	Sai phân bậc hai	Có	Có	Có
3	Sai phân mùa	Có	Có	Có
5	Trung bình trượt trung tâm, trọng số bằng nhau	Không	Không	Không
6	Thông thấp	Không	Có, không đáng kể	Không
7	Thông cao	Không	Có, không đáng kể	Có
8	Hodrick-Prescott	Không	Không	Có
9	Baxter-King	Không	Có, không đáng kể	Có

4.2.1. Phép lọc thông cao

4.2.1.1. Quy trình xây dựng phép lọc thông cao với tần số cắt $\lambda_0 = \frac{2\pi}{156}$

- *Bước 1:* Với mỗi giá trị K , xây dựng phép lọc thông thấp tần số cắt $\lambda_0 = \frac{2\pi}{156}$, ký hiệu là $LF_K(L) = \sum_{j=-K}^K b_j L^j$, trong đó, b_j được cho bởi công thức (2).

- *Bước 2:* Phép lọc thông cao, ký hiệu $HF_K(L)$ được xác định:

$$HF_K(L) = 1 - \sum_{j=-K}^K b_j L^j$$

- *Bước 3:* Lặp lại Bước 1 và Bước 2 với K tăng dần, cho đến khi độ lệch chuẩn và hệ số tự tương quan bậc 1 của các thành phần chu kỳ $C_t^K = HF_K(L)LVNI_t$ tương đối ổn định.

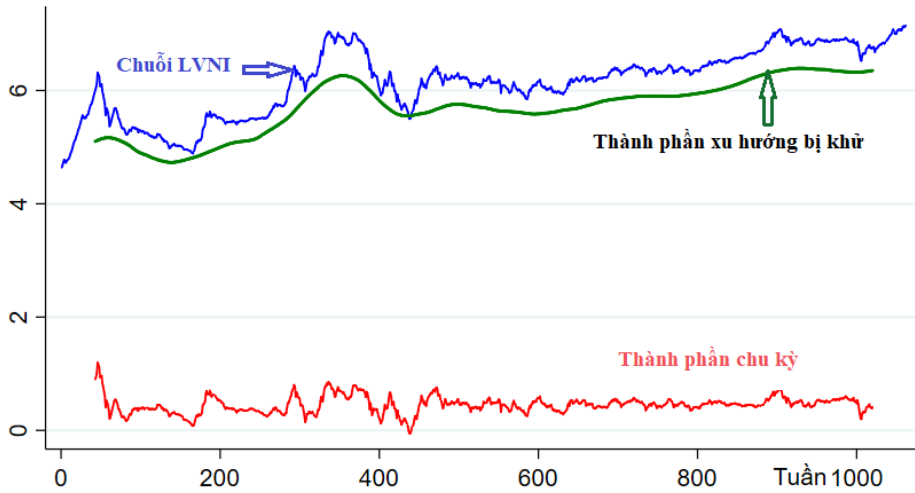
4.2.1.2. Độ ổn định và độ bền của các chu kỳ theo K

Bảng 5 trình bày các giá trị đo độ dao động và độ bền của phép lọc thông cao tương ứng với các giá trị của K . Từ $K = 4, \dots, 36$, độ lệch chuẩn và hệ số tự tương quan bậc 1 của các chu kỳ giảm dần. Từ $K = 42$ trở đi, các độ đo này dần ổn định. Vì thế, phép lọc thông cao nên được chọn khi $K \geq 42$. Ta biết rằng, khi K càng lớn, HF_K càng tiệm cận với phép lọc lý tưởng. Tuy nhiên, lúc đó chuỗi chu kỳ nhận được bị mất $2K$ quan sát. Vì thế, giá trị $K = 42$ được đề nghị trong phép lọc thông cao tần số cắt $\lambda_0 = \frac{2\pi}{156}$. Hình 3 thể hiện kết quả của phép lọc thông cao HF_{42} .

Bảng 5.

Độ dao động và độ bền của các chu kỳ tương ứng các phép lọc thông cao HF_K

K	4	8	12	16	20	24	30	36
Độ lệch chuẩn	0,5049	0,4428	0,3850	0,3325	0,2851	0,2438	0,1949	0,1629
Hệ số tự tương quan bậc 1	0,9971	0,9963	0,9951	0,9933	0,9909	0,9875	0,9801	0,9712
K	42	52	62	72	84	94	104	114
Độ lệch chuẩn	0,1464	0,1325	0,1412	0,1488	0,1502	0,1460	0,1407	0,1385
Hệ số tự tương quan bậc 1	0,9644	0,9588	0,9671	0,9703	0,9705	0,9683	0,9653	0,9636



Hình 3. Minh họa kết quả phép lọc thông cao HF_{42}

4.2.2. Phép lọc Hodrick-Prescott

4.2.2.1. Ước lượng hệ số làm trơn

Với λ cho trước, hàm lợi ích của HPF là: $G^{HP}(\omega) = \frac{(4\lambda(1-\cos\omega)^2)^2}{(4\lambda(1-\cos\omega)^2+1)^2}$

$$\text{Với } \omega \in \left(0; \frac{2\pi}{156}\right], \text{ ta có: } G^{HP}(\omega) \leq \frac{16\lambda^2(1-\cos\frac{2\pi}{156})^4}{(4\lambda(1-\cos\frac{2\pi}{156})^2+1)^2}$$

Nói cách khác, $G^{HP}(\omega) \leq \left(\frac{1}{1+\frac{1}{A\lambda}}\right)^2$ trong đó $A = 4\left(1 - \cos\frac{2\pi}{156}\right)^2 \approx 2,63 \times 10^{-6}$

Vì HPF là một phép lọc thông cao nên sẽ giữ lại các thành phần có chu kỳ không quá 156 tuần (loại bỏ các thành phần có tần số $\omega \in \left[0; \frac{2\pi}{156}\right]$). Do đó, λ được xác định thỏa $G^{HP}(\omega) \approx 0$ khi $\omega \in$

$\left(0; \frac{2\pi}{156}\right]$. Nếu cho trước $\varepsilon^2 \in (0,1)$, sao cho $\left(\frac{1}{1+\frac{1}{A\lambda}}\right)^2 = \varepsilon^2$ thì λ được tính theo công thức:

$$\lambda = \frac{\varepsilon}{A(1-\varepsilon)} \text{ với } A = 2,63 \times 10^{-6} \quad (9)$$

Bảng 6 giới thiệu một số giá trị của hệ số làm trơn λ tương ứng với các giá trị của ε . Nhận thấy, ε^2 càng lớn thì λ càng lớn, khi đó, HPF tiệm cận với phép lọc thông cao lý tưởng. Khi λ nhỏ (tương đương với ε^2 khá nhỏ), HPF không xấp xỉ tốt cho phép lọc thông cao lý tưởng. Tuy nhiên, λ quá lớn thì HPF sinh ra thành phần xu hướng khá phẳng, gần như tuyến tính. Vì thế, từ Bảng 6, các giá trị λ nên được chọn trong khoảng $[886.900; 1.520.400]$.

Bảng 6.

Ước lượng giá trị hệ số làm trơn của phép lọc Hodrick-Prescott

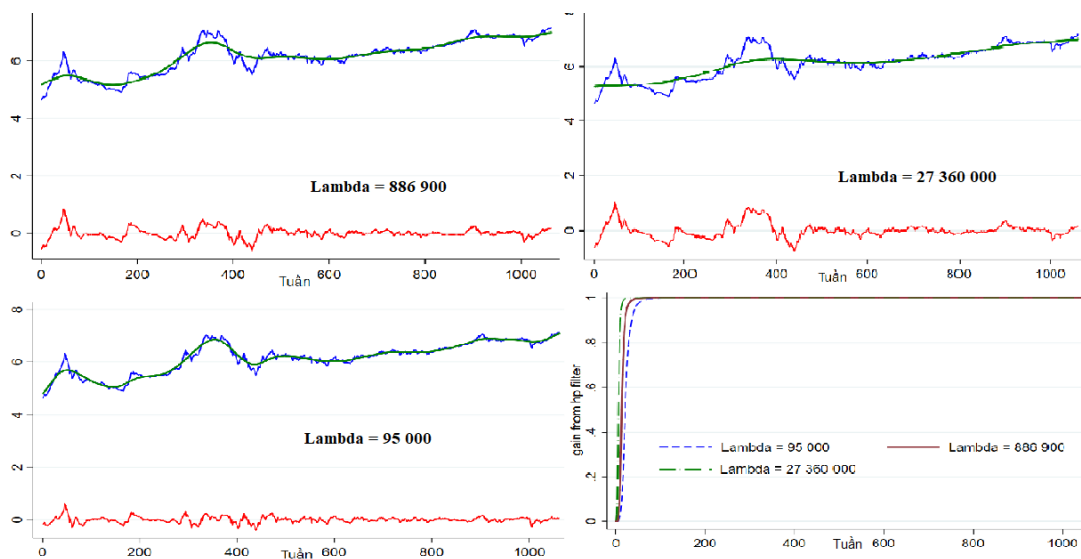
ε	ε^2	λ
0,80	0,64	1.520.400
0,70	0,49	886.900
0,60	0,36	570.150
0,50	0,25	380.100
0,40	0,16	253.400
0,30	0,09	162.900
0,20	0,04	95.000

4.2.2.2. Kết quả HFK với các giá trị λ

Phần này so sánh các chuỗi xu hướng nhận được từ các HPF tương ứng với các giá trị:

$$\lambda_U = 27.360.000; \lambda_1 = 886.900; \lambda_2 = 95.000$$

Trong đó, $\lambda_U = 27.360.000$ là giá trị được tính toán theo đề xuất của Uhlig cho bởi công thức (7) với $m = 3,8; \lambda_4 = 1.600$.



Hình 4. Minh họa một số phép lọc Hodrick-Prescott

So sánh với các giá trị λ ở Bảng 6, nhận thấy hệ số làm trơn theo đề xuất của Ravn và Uhlig (2002) là quá lớn. Minh họa ở Hình 4 cho thấy đường xu hướng của HPF tương ứng với λ_U khá phẳng. Như vậy, hệ số làm trơn theo đề xuất của Ravn và Uhlig (2002) không phù hợp chuỗi LVNI. Mặt khác, đường xu hướng tương ứng với $\lambda_2 = 95.000$ gần như theo sát với chuỗi LVNI. Tuy vậy, hàm lợi ích của phép lọc tương ứng lại là xấp xỉ kém nhất cho hàm lợi ích của phép lọc thông cao lý tưởng. Có thể cho rằng, trong các giá trị λ được xét, $\lambda_1 = 886.900$ là phù hợp nhất.

4.2.3. Phép lọc Baxter-King

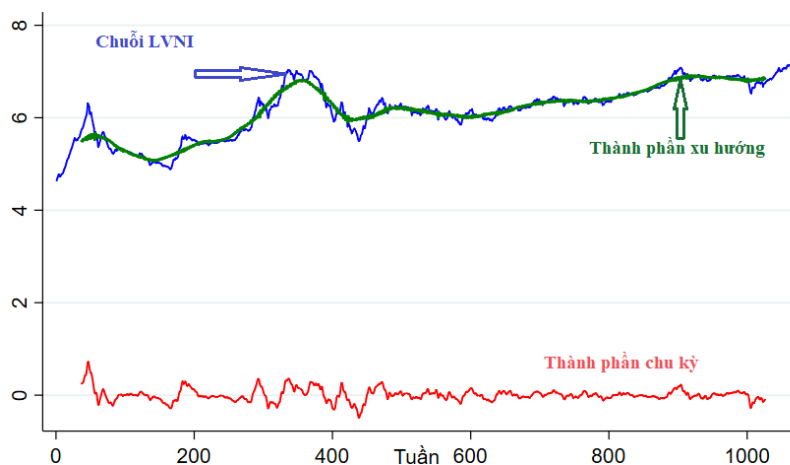
Chuỗi VN-Index có chu kỳ dao động trung bình là 156 tuần. Trong mỗi chu kỳ, thị trường chứng khoán có những dao động nhỏ điều chỉnh giá với thời gian không quá 4 tuần. Vì thế, phần này nghiên cứu các BKF với $p_1 = 4$ và $p_2 = 156$ tại những giá trị K khác nhau. Tương tự quy trình xây dựng phép lọc thông cao, độ lệch chuẩn và hệ số tự tương quan bậc 1 được dùng làm độ đo để xác định giá trị K hợp lý nhất cho BKF.

Bảng 7.

Độ dao động và độ bền của các $BKF_K(4, 156)$ theo K

K	4	8	12	16	20	24	30
Độ lệch chuẩn	0,0321	0,0515	0,0695	0,0842	0,0954	0,1063	0,1193
Hệ số tự tương quan bậc 1	0,7236	0,8705	0,9208	0,9402	0,9539	0,9620	0,9689
K	36	42	52	62	72	82	92
Độ lệch chuẩn	0,1319	0,1368	0,1347	0,1384	0,1421	0,1403	0,1383
Hệ số tự tương quan bậc 1	0,9739	0,9761	0,9768	0,9803	0,9815	0,9808	0,9800

Bảng 7 cho thấy khi $K \geq 36$, độ lệch chuẩn và hệ số tự tương quan bậc 1 của các chuỗi chu kỳ dần ổn định. Tương tự trường hợp xấp xỉ phép lọc thông cao, giữa yêu cầu xấp xỉ tốt phép lọc thông dài lý tưởng và sự mất mát dữ liệu mẫu, giá trị K phù hợp nhất là 36. Hình 5 minh họa các thành phần xu hướng và chu kỳ của chuỗi LVNI nhận được từ $BK_{36}(4, 156)$.



Hình 5. Minh họa kết quả phép lọc Baxter-King $BK_{36}(4, 156)$

4.2.4. So sánh các phép lọc thông cao, Hodrick-Prescott và Baxter-King

Như vậy, đã xác định được các phép lọc có thể tách thành phần xu hướng và chu kỳ của chuỗi LVNI, bao gồm:

- (1) HF với tần số cắt $\lambda_0 = \frac{2\pi}{156}$; $K = 42$.
- (2) HPF với hệ số làm trơn là $\lambda = 886.900$.
- (3) BKF với $p_1 = 4$; $p_2 = 156$; $K = 36$.

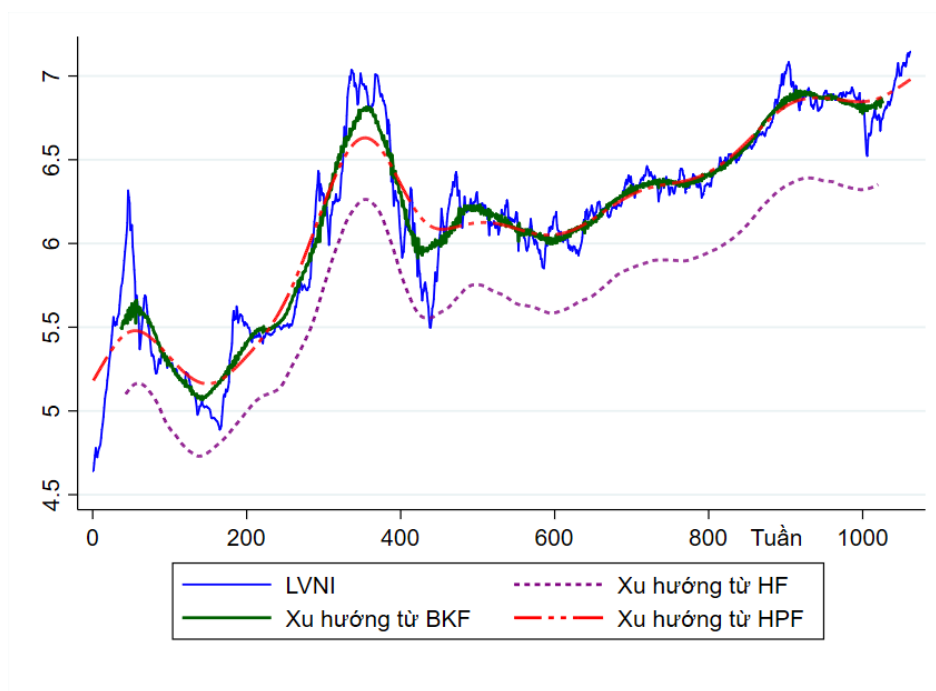
Bảng 8.

Ma trận hệ số tương quan của các đường xu hướng tương ứng HF, HPF, BKF

	LVNI	Trend-HF	Trend-HPF	Trend-BKF
LVNI	1,0000			
Trend-HF	0,9677	1,0000		
Trend-HPF	0,9520	0,9947	1,0000	
Trend-BKF	0,9711	0,9992	0,9917	1,0000

Ghi chú: Trend-HF42, Trend-HPF, Trend-BKF là thành phần xu hướng của chuỗi LVNI thu được từ các phép lọc HF, HPF và BKF.

Hình 6 biểu diễn các đường xu hướng của chuỗi LVNI được tách bởi ba phép lọc nói trên. Hệ số tương quan giữa chuỗi gốc LVNI và các đường xu hướng được cho ở Bảng 8. Các đường xu hướng khá tương đồng với chuỗi gốc, trong đó, hệ số tương quan giữa LVNI và đường xu hướng của BKF là cao nhất (0,9711).

**Hình 6.** Các đường xu hướng từ các phép lọc HF, HPF, BKF

Bên cạnh đó, đường xu hướng từ các phép lọc thông cao HF_{42} và Baxter-King BK_{36} rất giống nhau về xu hướng vận động. Vì thế, các chuỗi chu kỳ từ phép lọc HF_{42} và BK_{36} là tương tự nhau (hệ số tương quan trên cùng mẫu với 978 quan sát là (0,9531)). Mặt khác, phép lọc Baxter-King BK_{36} ít làm mất dữ liệu hơn so với phép lọc thông cao HF_{42} . Vì thế, BK_{36} là lựa chọn tốt hơn so với HF_{42} .

Phép lọc Hodrick-Prescott sinh ra chuỗi chu kỳ với đầy đủ quan sát của chuỗi gốc. Tuy vậy, hệ số tương quan của đường xu hướng sinh bởi HPF và chuỗi LVNI là nhỏ nhất trong ba phép lọc được khảo sát (0,9520). Mặt khác, ở các thời điểm đầu và cuối dữ liệu, đường xu hướng có vẻ không khớp

với chuỗi dữ liệu gốc. Điều này tương đồng với các nhận định của các nghiên cứu thực nghiệm trước đây (Baxter & King, 1999; Hamilton, 2018). Để tránh vấn đề giả mạo thông tin ở chuỗi đầu ra của HPF, cần loại dữ liệu ở hai đầu chuỗi. Nếu căn cứ vào kích thước mẫu của BK_{36} , số quan sát nên bị loại bỏ là 72 (36 quan sát ở đầu chuỗi và 36 quan sát ở cuối chuỗi).

Như vậy, đối với chuỗi LVNI, phép lọc Baxter-King $BK_{36}(4, 156)$ là lựa chọn tốt nhất để tách thành phần xu hướng. $BK_{36}(4, 15)$ đồng thời tách được thành phần bất quy tắc, có chu kỳ dao động không quá 4 tuần.

5. Kết luận

Bài báo đánh giá các tính chất đặc trưng của các phép lọc tuyến tính phổ biến trong phân tích chuỗi thời gian thông qua nghiên cứu hàm truyền và hàm lợi ích. Các phép sai phân bậc nhất, sai phân bậc hai và sai phân mùa tuy có thể loại bỏ thành phần xu hướng, nhưng lại phóng đại biên độ dao động của một số thành phần, đồng thời làm thay đổi mối quan hệ về thời gian của các sự kiện trong chuỗi chu kỳ so với chuỗi gốc. Phép trung bình trượt trung tâm là phép lọc thông thấp nên không thể tách được thành phần có chu kỳ dài. Vì thế, với mục đích loại bỏ thành phần xu hướng, không làm thay đổi pha của chuỗi gốc và không phóng đại biên độ dao động của chuỗi chu kỳ, các phép lọc thông cao, Hodrick-Prescott và Baxter-King là những phép lọc có thể được sử dụng.

Trên chuỗi logarit tự nhiên của chỉ số VN-Index tần số tuần, để loại bỏ xu hướng (là thành phần có chu kỳ lớn hơn 156 tuần), nhóm tác giả đã tìm được các tham số hợp lý nhất cho từng phép lọc. Đó là phép lọc thông cao HF_{42} tần số cắt $\frac{2\pi}{156}$, phép lọc Hodrick-Prescott với tham số làm trơn $\lambda = 886.900$, và phép lọc Baxter-King $BK_{36}(4, 156)$. Hơn nữa, phép lọc Baxter-King loại bỏ cả thành phần có chu kỳ không quá 4 tuần. Trong đó, phép lọc Baxter-King cho kết quả tốt nhất. Các đề nghị tham số của các phép lọc này có thể được áp dụng trên chuỗi dữ liệu tuần, có chu kỳ dao động là 156 tuần tại Việt Nam.

Tài liệu tham khảo

- Baxter, M., & King, R. G. (1999). Measuring business cycles: Approximate band-pass filters for economic time series. *Review of Economics and Statistics*, 81(4), 575–593.
- Dadashova, B. (2012). *Detrending the business cycles: Hodrick-Prescott and Baxter-King filters*, 1–23. Universidad Carlos III de Madrid.
- Hamilton, J. D. (2020). *Time Series Analysis*. United States: Princeton University Press.
- Hamilton, J. D. (2018). Why you should never use the Hodrick-Prescott filter. *Review of Economics and Statistics*, 100(5), 831–843. doi: 10.1162/rest_a_00706
- Hodrick, R., & Prescott, E. (1997). Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1), 1–16. doi: 10.2307/2953682
- Hornsterin, A. (1998). Inventory investment and the business cycle. *Federal Reserve Bank Richmond Economic Quarterly*, 84(2), 47–71.

- Iacobucci, A. (2005). Spectral analysis for economic time series. In Leskow J., Punzo L. F., Anyul M. P. (eds), *Lecture Notes in Economics and Mathematical Systems*, 551, 203–219. Berlin, Heidelberg: Springer. doi: 10.1007/3-540-28444-3_12
- Kamalian, A., Zamani, Z., Amirali, M., & Dehkordi, M. M. (2020). Analyzing the different effects of endogenous and exogenous money supply on inflation: A spectral analysis approach. *The Economic Research*, 20(3), 57–77. Retrieved from <https://ecor.modares.ac.ir/article-18-37417-en.html>
- Neusser, K. (2016). *Time Series Econometrics*. Berlin, Heidelberg: Springer.
- Ravn, M. O., & Uhlig, H. (2002). On adjusting the Hodrick-Prescott filter for the frequency of observations. *Review of Economics and Statistics*, 84(2), 371–376. doi: 10.1162/003465302317411604
- Tastan, H., & Sahin, S. (2020). Low-frequency relationship between money growth and inflation in Turkey. *Quantitative Finance and Economics*, 4(1), 91–120. doi: 10.3934/qfe.2020005