

Boosted HP Filter: Several Properties Derived from Its Spectral Representation

Hiroshi Yamada(✉)

Hiroshima University, Higashi-Hiroshima 739-8525, Japan
yamada@hiroshima-u.ac.jp

Abstract. The Hodrick–Prescott (HP) filter is the most prominent smoothing/trend estimation method for macroeconomic time series such as real gross domestic product. Recently, its modification, the boosted HP (bHP) filter, has been developed, and a number of studies have been conducted on it. This paper contributes to the literature by investigating the properties of the bHP filter. For this purpose, we use the spectral decomposition of the penalty matrix of the HP filter.

Keywords: Hodrick–Prescott filter · boosted HP filter · Whittaker–Henderson graduation

1 Introduction

The Hodrick–Prescott (HP) filter [1] is the most prominent smoothing/trend estimation method for macroeconomic time series such as real gross domestic product. As is well known, it is a kind of Whittaker–Henderson (WH) graduation, which is frequently used in the actuarial literature [2]. Recently, Phillips and Shi [3] developed its modification, the boosted HP (bHP) filter, and a number of studies concerning the filter have emerged [4–10]. In this paper, we contribute to the literature by investigating the properties of the bHP filter. For this purpose, we use the spectral decomposition of the penalty matrix of the HP filter. (The matrix is a symmetric pentadiagonal matrix as shown in the Appendix.) This is because, as we will show later, it allows us to derive the spectral decomposition of the smoother matrix of the bHP filter.

The organization of the paper is as follows. In Sect. 2, we provide some preliminary remarks, which include the spectral decomposition of the penalty matrix of the HP filter. In Sect. 3, we present the spectral representation of the bHP filter. In Sect. 4, we document several properties of the bHP filter. Section 5 concludes the paper. The Appendix contains some proofs.

2 Preliminaries

In this section, we provide some preliminary remarks.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
O. Gervasi et al. (Eds.): ICCSA 2024, LNCS 14814, pp. 230–242, 2024.
https://doi.org/10.1007/978-3-031-64608-9_15

Data. Let y_i denote the realization of a variable y at time i for $i = 1, \ldots, n$. In the paper, we assume that y_i cannot be represented as $\phi_1 + \phi_2 i$ for $i = 1, \ldots, n$, where both ϕ_1 and ϕ_2 are real numbers. This is because there is no need for smoothing in the case.

Some of the Notations. Let $\boldsymbol{y} = [y_1, \ldots, y_n]^\top$, $\boldsymbol{x} = [x_1, \ldots, x_n]^\top$, $\boldsymbol{I}_n$ be the $n \times n$ identity matrix, $\boldsymbol{\iota}$ be the n-dimensional column vector of ones, i.e., $\boldsymbol{\iota} = [1, \ldots, 1]^\top$, $\boldsymbol{Q}_\iota = \boldsymbol{I}_n - \boldsymbol{\iota}(\boldsymbol{\iota}^\top\boldsymbol{\iota})^{-1}\boldsymbol{\iota}^\top = \boldsymbol{I}_n - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top$, $\boldsymbol{\tau} = [1, \ldots, n]^\top$, and

$$\boldsymbol{\Pi} = [\boldsymbol{\iota}, \boldsymbol{\tau}].$$

In addition, let $\boldsymbol{\Delta}_2$ be the $(n-2) \times n$ second differencing matrix such that $\boldsymbol{\Delta}_2\boldsymbol{x} = (x_3 - 2x_2 + x_1, \ldots, x_n - 2x_{n-1} + x_{n-2})^\top$. Explicitly, $\boldsymbol{\Delta}_2$ is the Toeplitz matrix given by

$$\boldsymbol{\Delta}_2 = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & -2 & 1 \end{bmatrix}.$$

For a column vector $\boldsymbol{\eta}$, $\|\boldsymbol{\eta}\|^2 = \boldsymbol{\eta}^\top\boldsymbol{\eta}$. For a full column rank matrix $\boldsymbol{R}$, denote the column space of $\boldsymbol{R}$ by $\mathbb{S}(\boldsymbol{R})$ and let $\boldsymbol{P}_R = \boldsymbol{R}(\boldsymbol{R}^\top\boldsymbol{R})^{-1}\boldsymbol{R}^\top$. Accordingly, $\boldsymbol{P}_R$ is the orthogonal projection matrix onto $\mathbb{S}(\boldsymbol{R})$ and $\boldsymbol{Q}_\iota$ can be represented as $\boldsymbol{I}_n - \boldsymbol{P}_\iota$.

HP Filter. The HP filter is defined by the following minimization problem:

$$\min_{x_1,\ldots,x_n} f(x_1, \ldots, x_n) = \sum_{i=1}^{n}(y_i - x_i)^2 + \lambda \sum_{i=3}^{n}(x_i - 2x_{i-1} + x_{i-2})^2, \tag{1}$$

where $\lambda \in (0, \infty)$ is a smoothing parameter that controls fidelity, $\sum_{i=1}^{n}(y_i - x_i)^2$, and smoothness, $\sum_{i=3}^{n}(x_i - 2x_{i-1} + x_{i-2})^2$. (1) can be represented in matrix form as

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \boldsymbol{x}^\top \boldsymbol{C}\boldsymbol{x}, \tag{2}$$

where

$$\boldsymbol{C} = \boldsymbol{\Delta}_2^\top \boldsymbol{\Delta}_2. \tag{3}$$

Accordingly, $\boldsymbol{C}$ is the penalty matrix of the HP filter. Given that $f(\boldsymbol{x})$ is a quadratic function of $\boldsymbol{x}$ whose Hessian matrix, $2(\boldsymbol{I}_n + \lambda\boldsymbol{C})$, is positive definite, there exists $\widehat{\boldsymbol{x}}^{\mathrm{HP}}$ such that $f(\boldsymbol{x}) > f(\widehat{\boldsymbol{x}}^{\mathrm{HP}})$ for all $\boldsymbol{x} \in \mathbb{R}^n \backslash \{\widehat{\boldsymbol{x}}^{\mathrm{HP}}\}$. Explicitly, it is given by the following linear transformation of $\boldsymbol{y}$,

$$\widehat{\boldsymbol{x}}^{\mathrm{HP}} = \boldsymbol{A}\boldsymbol{y}, \tag{4}$$

where

$$\boldsymbol{A} = (\boldsymbol{I}_n + \lambda \boldsymbol{C})^{-1}. \tag{5}$$

Note that $(\boldsymbol{I}_n + \lambda \boldsymbol{C})$ in (5) is nonsingular because it is positive definite [11]. Thus, the HP filter is a linear smoother and $\boldsymbol{A}$ is referred to as the smoother matrix of it. In addition, $\boldsymbol{C}$ in $\boldsymbol{A}$ is referred to as the penalty matrix of the HP filter.

bHP Filter. The trend estimated by the bHP filtering is given by the following linear transformation of $\boldsymbol{y}$,

$$\widehat{\boldsymbol{x}}^{(m)} = \boldsymbol{A}_m \boldsymbol{y}, \tag{6}$$

where

$$\boldsymbol{A}_m = \boldsymbol{I}_n - (\boldsymbol{I}_n - \boldsymbol{A})^m. \tag{7}$$

Thus, the bHP filter is also a linear smoother. We refer to $\boldsymbol{A}_m$ as the smoother matrix of the bHP filter. In addition, we refer to m as the degree of boosting. Then, given that

$$\boldsymbol{A}_1 = \boldsymbol{I}_n - (\boldsymbol{I}_n - \boldsymbol{A})^1 = \boldsymbol{A}, \tag{8}$$

it follows that

$$\widehat{\boldsymbol{x}}^{(1)} = \boldsymbol{A}_1 \boldsymbol{y} = \boldsymbol{A}\boldsymbol{y} = \widehat{\boldsymbol{x}}^{\mathrm{HP}}. \tag{9}$$

Thus, the bHP filter is a generalization of the HP filter.

Let $\boldsymbol{A}_m = [a_{i,j}^{(m)}]$ and $\widehat{x}_i^{(m)}$ denote the i-th entry of $\widehat{\boldsymbol{x}}^{(m)}$. Then, (6) can be represented as

$$\widehat{x}_i^{(m)} = a_{i,1}^{(m)} y_1 + \cdots + a_{i,n}^{(m)} y_n. \tag{10}$$

Spectral Decomposition of $\boldsymbol{C}$. Since $\boldsymbol{C}$ in (3) is a real symmetric matrix, it can be spectrally decomposed as

$$\boldsymbol{C} = \boldsymbol{V}\boldsymbol{D}\boldsymbol{V}^\top, \tag{11}$$

where $\boldsymbol{D} = \mathrm{diag}(d_1, \ldots, d_n)$ and $\boldsymbol{V} = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]$ is an orthogonal matrix. Here, $d_1, \ldots, d_n$ are in ascending order. Since $\boldsymbol{C}$ is a positive semidefinite matrix whose rank equals $n-2$, it follows that $d_1 = d_2 = 0$ and $d_i > 0$ for $i = 3, \ldots, n$. With regard to $d_1 = d_2 = 0$, given $\boldsymbol{C}\boldsymbol{\iota} = 0 \cdot \boldsymbol{\iota}$, $\boldsymbol{C}(\boldsymbol{Q}_\iota \boldsymbol{\tau}) = 0 \cdot (\boldsymbol{Q}_\iota \boldsymbol{\tau})$, and $\boldsymbol{\iota}^\top (\boldsymbol{Q}_\iota \boldsymbol{\tau}) = 0$, we can let $(d_1, \boldsymbol{v}_1) = (0, \boldsymbol{\iota}/\sqrt{n})$ and $(d_2, \boldsymbol{v}_2) = (0, \boldsymbol{Q}_\iota \boldsymbol{\tau}/\sqrt{\boldsymbol{\tau}^\top \boldsymbol{Q}_\iota \boldsymbol{\tau}})$. With regard to d_n, which denotes the largest eigenvalue of $\boldsymbol{C}$, it follows that

$$d_n \leq 16. \tag{12}$$

For example, when $n = 100$, $d_n = 15.9920$. For proofs of (12), see the Appendix. In summary, the following inequalities apply to the eigenvalues of $\boldsymbol{C}$, $d_1, \ldots, d_n$:

$$0 = d_1 = d_2 < d_3 \leq \cdots \leq d_n \leq 16, \tag{13}$$

which is used in the next section.

Spectral Decomposition of $\boldsymbol{A}$. The smoother matrix of the HP filter, $\boldsymbol{A}$, which is a real symmetric matrix, can be spectrally decomposed as

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{B}\boldsymbol{V}^\top, \tag{14}$$

where $\boldsymbol{B} = (\boldsymbol{I}_n + \lambda \boldsymbol{D})^{-1}$. Let

$$b_i = \frac{1}{1 + \lambda d_i}, \quad i = 1, \ldots, n. \tag{15}$$

Then, $\boldsymbol{B} = \mathrm{diag}(b_1, \ldots, b_n)$.

Spectral Decomposition of $\boldsymbol{A}_m$. The smoother matrix of the bHP filter, $\boldsymbol{A}_m$, which is a real symmetric matrix, can be spectrally decomposed as

$$\boldsymbol{A}_m = \boldsymbol{V}\boldsymbol{B}_m\boldsymbol{V}^\top, \tag{16}$$

where $\boldsymbol{B}_m = \boldsymbol{I}_n - (\boldsymbol{I}_n - \boldsymbol{B})^m$. (16) can be derived from (14) as

$$\begin{aligned} \boldsymbol{A}_m &= \boldsymbol{I}_n - (\boldsymbol{I}_n - \boldsymbol{A})^m = \boldsymbol{I}_n - (\boldsymbol{I}_n - \boldsymbol{V}\boldsymbol{B}\boldsymbol{V}^\top)^m \\ &= \boldsymbol{V}\{\boldsymbol{I}_n - (\boldsymbol{I}_n - \boldsymbol{B})^m\}\boldsymbol{V}^\top = \boldsymbol{V}\boldsymbol{B}_m\boldsymbol{V}^\top. \end{aligned}$$

Then, it follows that

$$\boldsymbol{B}_1 = \boldsymbol{I}_n - (\boldsymbol{I}_n - \boldsymbol{B})^1 = \boldsymbol{B}. \tag{17}$$

This is consistent with (8). It should be noted that (16) is a restatement of the result in [4].

Let

$$b_i^{(m)} = 1 - \left(\frac{\lambda d_i}{1 + \lambda d_i}\right)^m, \quad i = 1, \ldots, n. \tag{18}$$

Then, $\boldsymbol{B}_m = \mathrm{diag}(b_1^{(m)}, \ldots, b_n^{(m)})$. From (15) and (18), it follows that

$$b_i^{(1)} = 1 - \left(\frac{\lambda d_i}{1 + \lambda d_i}\right)^1 = \frac{1 + \lambda d_i - \lambda d_i}{1 + \lambda d_i} = \frac{1}{1 + \lambda d_i} = b_i, \quad i = 1, \ldots, n. \tag{19}$$

3 Spectral Representation of the bHP Filter

From (16), which is a spectral decomposition of $\boldsymbol{A}_m$, the bHP filter can be represented as

$$\widehat{\boldsymbol{x}}^{(m)} = \boldsymbol{V}\boldsymbol{B}_m\boldsymbol{V}^\top\boldsymbol{y} \tag{20}$$

$$= b_1^{(m)} z_1 \boldsymbol{v}_1 + \cdots + b_n^{(m)} z_n \boldsymbol{v}_n, \tag{21}$$

where $\boldsymbol{z} = [z_1, \ldots, z_n]^\top = \boldsymbol{V}^\top \boldsymbol{y}$. We will refer to (20)/(21) as the spectral representation of the bHP filter. It is the linear combination of an orthonormal basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ such that the coefficient of $\boldsymbol{v}_i$ is $b_i^{(m)} z_i$ for $i = 1, \ldots, n$.

We make four points regarding the spectral representation of the bHP filter.

1. Given that

$$\|\boldsymbol{\Delta}_2 \boldsymbol{v}_i\|^2 = \boldsymbol{v}_i^\top \boldsymbol{C} \boldsymbol{v}_i = d_i, \quad i = 1, \ldots, n,$$

from (13), the following inequalities hold.

$$0 = \|\boldsymbol{\Delta}_2 \boldsymbol{v}_1\|^2 = \|\boldsymbol{\Delta}_2 \boldsymbol{v}_2\|^2 < \|\boldsymbol{\Delta}_2 \boldsymbol{v}_3\|^2 \leq \cdots \leq \|\boldsymbol{\Delta}_2 \boldsymbol{v}_n\|^2 \leq 16. \tag{22}$$

These indicate that with respect to $\boldsymbol{\Delta}_2$,
(a) both $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are perfectly smooth,
(b) $\boldsymbol{v}_i$ is smoother than $\boldsymbol{v}_{i+1}$ for $i = 2, \ldots, n-1$, and
(c) $\|\boldsymbol{\Delta}_2 \boldsymbol{v}_i\|^2$ is at most 16.
See Fig. 1, which depicts $\boldsymbol{v}_3, \ldots, \boldsymbol{v}_{11}$ when $n = 100$.

2. It follows that

$$1 = b_1^{(m)} = b_2^{(m)} > b_3^{(m)} \geq \cdots \geq b_n^{(m)} > 0. \tag{23}$$

Therefore, the contribution of eigenvectors corresponding to larger eigenvalues becomes smaller. (In addition, from (23), $\boldsymbol{A}_m$ is positive definite.)

3. It follows that

$$z_i = \boldsymbol{v}_i^\top \boldsymbol{y} = (\boldsymbol{v}_i^\top \boldsymbol{v}_i)^{-1} \boldsymbol{v}_i^\top \boldsymbol{y} = \arg\min_{\zeta} \|\boldsymbol{y} - \zeta \boldsymbol{v}_i\|^2, \quad i = 1, \ldots, n. \tag{24}$$

Moreover, denoting the sample correlation coefficient between $\boldsymbol{v}_i$ and $\boldsymbol{y}$ by $\mathrm{cor}(\boldsymbol{v}_i, \boldsymbol{y})$, it follows that

$$\mathrm{cor}(\boldsymbol{v}_i, \boldsymbol{y}) = \frac{\boldsymbol{v}_i^\top \boldsymbol{Q}_\iota \boldsymbol{y}}{\sqrt{\boldsymbol{v}_i^\top \boldsymbol{Q}_\iota \boldsymbol{v}_i} \sqrt{\boldsymbol{y}^\top \boldsymbol{Q}_\iota \boldsymbol{y}}} = \frac{z_i}{\sqrt{\boldsymbol{y}^\top \boldsymbol{Q}_\iota \boldsymbol{y}}}, \quad i = 2, \ldots, n. \tag{25}$$

Here, the second equality follows from $\boldsymbol{v}_i^\top \boldsymbol{v}_i = 1$ and $\boldsymbol{v}_i^\top \boldsymbol{\iota} = 0$ for $i = 2, \ldots, n$. (Note that since $\boldsymbol{y} \notin \mathbb{S}(\boldsymbol{\Pi})$ by assumption, $\boldsymbol{y}^\top \boldsymbol{Q}_\iota \boldsymbol{y}$ in (25) is positive.) Therefore, given $\boldsymbol{y}$, z_i is proportional to $\mathrm{cor}(\boldsymbol{v}_i, \boldsymbol{y})$.

4. Letting

$$\widehat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}} \|\boldsymbol{y} - \boldsymbol{\Pi}\boldsymbol{\gamma}\|^2 = (\boldsymbol{\Pi}^\top \boldsymbol{\Pi})^{-1} \boldsymbol{\Pi}^\top \boldsymbol{y},$$

it follows that

$$b_1^{(m)} z_1 \boldsymbol{v}_1 + b_2^{(m)} z_2 \boldsymbol{v}_2 = \boldsymbol{\Pi} \widehat{\boldsymbol{\gamma}} = \boldsymbol{P}_{\boldsymbol{\Pi}} \boldsymbol{y}, \tag{26}$$

which is the orthogonal projection of $\boldsymbol{y}$ onto $\mathbb{S}(\boldsymbol{\Pi})$. For a proof of (26), see the Appendix.

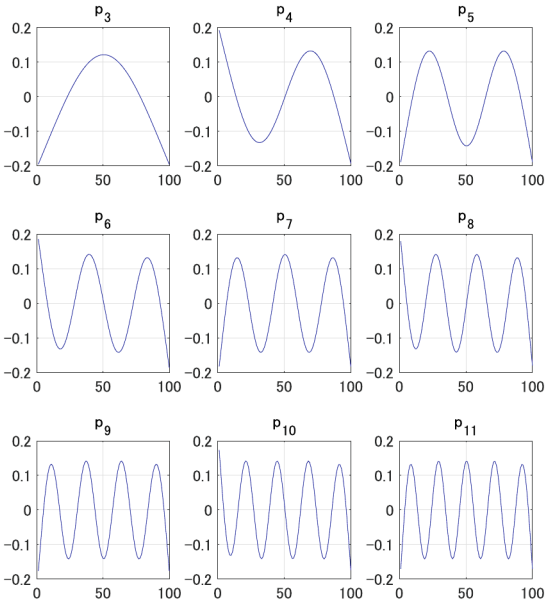

Fig. 1. $\boldsymbol{v}_i$ such that $\boldsymbol{C}\boldsymbol{v}_i = d_i\boldsymbol{v}_i$ and $\boldsymbol{v}_i^\top\boldsymbol{v}_i = 1$ for $i = 3, \ldots, 11$ when $n = 100$.

4 Several Properties of the bHP Filter

The spectral decomposition of the smoother matrix of the bHP filter given in (16) and the spectral representation of the bHP filter given by (20)/(21) allow us to derive several properties of the bHP filter. In this section, we describe such properties. Those are summarized in Proposition 1.

1. Given that $b_1^{(m)} = 1$ and $z_1 = \boldsymbol{v}_1^\top\boldsymbol{y}$, premultiplying (21) by $\boldsymbol{v}_1^\top$ yields

$$\boldsymbol{v}_1^\top\widehat{\boldsymbol{x}}^{(m)} = b_1^{(m)}z_1\boldsymbol{v}_1^\top\boldsymbol{v}_1 = \boldsymbol{v}_1^\top\boldsymbol{y}. \tag{27}$$

Since $\boldsymbol{v}_1 = \boldsymbol{\iota}/\sqrt{n}$, multiplying (27) by $\sqrt{n}$ yields

$$\boldsymbol{\iota}^\top\widehat{\boldsymbol{x}}^{(m)} = \boldsymbol{\iota}^\top\boldsymbol{y}. \tag{28}$$

Thus, the mean of the entries of $\widehat{\boldsymbol{x}}^{(m)}$ equals that of $\boldsymbol{y}$. That is, it follows that

$$\frac{1}{n}\sum_{i=1}^{n}\widehat{x}_i^{(m)} = \frac{1}{n}\sum_{i=1}^{n}y_i. \tag{29}$$

In addition, from (28), it immediately follows that

$$\boldsymbol{\iota}^\top(\boldsymbol{y} - \widehat{\boldsymbol{x}}^{(m)}) = 0. \tag{30}$$

Thus, the bHP filter trend residuals sum to zero. That is, it follows that

$$\sum_{i=1}^{n}(y_i - \widehat{x}_i^{(m)}) = 0. \tag{31}$$

2. Given that $(b_1^{(m)}, \boldsymbol{v}_1)$ is an eigenpair of $\boldsymbol{A}_m$ and $b_1^{(m)} = 1$, postmultiplying (16) by $\boldsymbol{v}_1$ yields

$$\boldsymbol{A}_m \boldsymbol{v}_1 = b_1^{(m)} \boldsymbol{v}_1 = \boldsymbol{v}_1. \tag{32}$$

Again, since $\boldsymbol{v}_1 = \boldsymbol{\iota}/\sqrt{n}$, from (32), it follows that

$$\boldsymbol{A}_m \boldsymbol{\iota} = \boldsymbol{\iota}. \tag{33}$$

Thus, each row of the smoother matrix $\boldsymbol{A}_m$ sums to unity. That is, it follows that

$$a_{i,1}^{(m)} + \cdots + a_{i,n}^{(m)} = 1, \quad i = 1, \ldots, n, \tag{34}$$

in (10).

3. Similarly, given that $(b_2^{(m)}, \boldsymbol{v}_2)$ is an eigenpair of $\boldsymbol{A}_m$ and $b_2^{(m)} = 1$, postmultiplying (16) by $\boldsymbol{v}_2$ yields

$$\boldsymbol{A}_m \boldsymbol{v}_2 = b_2^{(m)} \boldsymbol{v}_2 = \boldsymbol{v}_2. \tag{35}$$

Since $\boldsymbol{v}_2 = \boldsymbol{Q}_\iota \boldsymbol{\tau}/\sqrt{\boldsymbol{\tau}^\top \boldsymbol{Q}_\iota \boldsymbol{\tau}}$, from (35), we obtain

$$\boldsymbol{A}_m(\boldsymbol{Q}_\iota \boldsymbol{\tau}) = (\boldsymbol{Q}_\iota \boldsymbol{\tau}). \tag{36}$$

Then, from (33), it follows that $\boldsymbol{A}_m \boldsymbol{Q}_\iota = \boldsymbol{A}_m - \boldsymbol{P}_\iota$. Thus, (36) becomes $\boldsymbol{A}_m \boldsymbol{\tau} - \boldsymbol{P}_\iota \boldsymbol{\tau} = \boldsymbol{\tau} - \boldsymbol{P}_\iota \boldsymbol{\tau}$, which leads to

$$\boldsymbol{A}_m \boldsymbol{\tau} = \boldsymbol{\tau}. \tag{37}$$

Combining (33) and (37) yields

$$\boldsymbol{A}_m \boldsymbol{\Pi} = \boldsymbol{\Pi}. \tag{38}$$

Let $\boldsymbol{\xi} \in \mathbb{S}(\boldsymbol{\Pi})$. Then, from (38), it follows that which leads to $\boldsymbol{A}_m \boldsymbol{\xi} = \boldsymbol{\xi}$. Accordingly, since $\boldsymbol{P}_\Pi \boldsymbol{y} \in \mathbb{S}(\boldsymbol{\Pi})$, we obtain

$$\widehat{\boldsymbol{x}}^{(m)} = \boldsymbol{P}_\Pi \boldsymbol{y} + \boldsymbol{A}_m(\boldsymbol{y} - \boldsymbol{P}_\Pi \boldsymbol{y}). \tag{39}$$

4. Since

$$b_i^{(m)} = 1 - \left(\frac{\lambda d_i}{1 + \lambda d_i}\right)^m \to 0, \quad (n, m : \text{fixed}, \ \lambda \to \infty)$$

for $i = 3, \ldots, n$, it follows that

$$\begin{aligned} \widehat{\boldsymbol{x}}^{(m)} &= \boldsymbol{P}_\Pi \boldsymbol{y} + b_3^{(m)} z_3 \boldsymbol{v}_3 + \cdots + b_n^{(m)} z_n \boldsymbol{v}_n \\ &\to \boldsymbol{P}_\Pi \boldsymbol{y}, \quad (n, m : \text{fixed}, \ \lambda \to \infty). \end{aligned} \tag{40}$$

5. Since $b_1^{(m)} = b_2^{(m)} = 1$ and

$$b_i^{(m)} = 1 - \left(\frac{\lambda d_i}{1+\lambda d_i}\right)^m \to 1, \quad (n, m : \text{fixed}, \ \lambda \to 0)$$

for $i = 3, \ldots, n$, it follows that

$$\boldsymbol{A}_m = \boldsymbol{V}\boldsymbol{B}_m\boldsymbol{V}^\top \to \boldsymbol{V}\boldsymbol{V}^\top = \boldsymbol{I}_n, \quad (n, m : \text{fixed}, \ \lambda \to 0).$$

Therefore, it follows that

$$\widehat{\boldsymbol{x}}^{(m)} = \boldsymbol{A}_m\boldsymbol{y} \to \boldsymbol{y}, \quad (n, m : \text{fixed}, \ \lambda \to 0). \tag{41}$$

6. Since $b_1^{(m)} = b_2^{(m)} = 1$ and

$$b_i^{(m)} = 1 - \left(\frac{\lambda d_i}{1+\lambda d_i}\right)^m \to 1, \quad (n, \lambda : \text{fixed}, \ m \to \infty) \tag{42}$$

for $i = 3, \ldots, n$, it follows that

$$\boldsymbol{A}_m = \boldsymbol{V}\boldsymbol{B}_m\boldsymbol{V}^\top \to \boldsymbol{V}\boldsymbol{V}^\top = \boldsymbol{I}_n, \quad (n, \lambda : \text{fixed}, \ m \to \infty). \tag{43}$$

See Fig. 2. It illustrates (43) when $n = 100$ and $\lambda = 1600$. (Recall that each row of the smoother matrix $\boldsymbol{A}_m$ sums to unity.) Therefore, it follows that

$$\widehat{\boldsymbol{x}}^{(m)} = \boldsymbol{A}_m\boldsymbol{y} \to \boldsymbol{y}, \quad (n, \lambda : \text{fixed}, \ m \to \infty). \tag{44}$$

7. Given $\boldsymbol{A}_m^k = \boldsymbol{V}\boldsymbol{B}_m^k\boldsymbol{V}^\top$, (23), and (26), it follows that

$$\boldsymbol{A}_m^k\boldsymbol{y} = \boldsymbol{V}\boldsymbol{B}_m^k\boldsymbol{V}^\top\boldsymbol{y} \to \boldsymbol{V}_1\boldsymbol{V}_1^\top\boldsymbol{y} = \boldsymbol{P}_{\Pi}\boldsymbol{y}, \quad (n, m, \lambda : \text{fixed}, \ k \to \infty), \tag{45}$$

where $\boldsymbol{V}_1 = [\boldsymbol{v}_1, \boldsymbol{v}_2]$.

8. Since $\boldsymbol{A}_m$ is positive definite and λ is positive, we can let

$$\boldsymbol{C}_m = \lambda^{-1}(\boldsymbol{A}_m^{-1} - \boldsymbol{I}_n). \tag{46}$$

Then, $\boldsymbol{A}_m$ can be represented by $\boldsymbol{C}_m$ as

$$\boldsymbol{A}_m = (\boldsymbol{I}_n + \lambda\boldsymbol{C}_m)^{-1}, \tag{47}$$

which is the Reinsch form representation of $\boldsymbol{A}_m$. Accordingly, $\widehat{\boldsymbol{x}}^{(m)}$ is the solution of the following penalized least squares problem:

$$\min_{\boldsymbol{x}} f^{(m)}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda\boldsymbol{x}^\top\boldsymbol{C}_m\boldsymbol{x}, \tag{48}$$

and therefore $\boldsymbol{C}_m$ is the penalty matrix of the bHP filter.

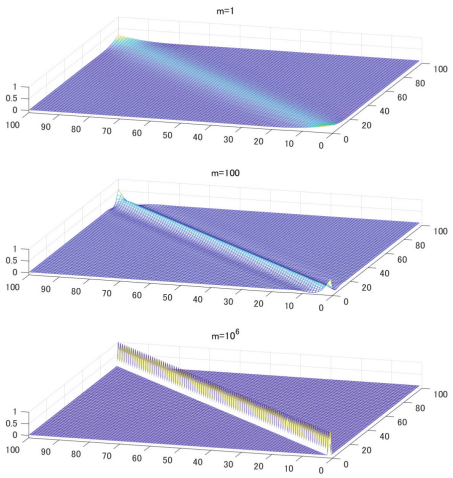

Fig. 2. Smoother matrix $\boldsymbol{A}_m$ in (7) for $m = 1, 100, 10^6$ when $n = 100$ and $\lambda = 1600$.

9. Given $\boldsymbol{A}_m = \boldsymbol{V}\boldsymbol{B}_m\boldsymbol{V}^\top$, it follows that $\boldsymbol{I}_n + \lambda\boldsymbol{C}_m = \boldsymbol{V}\boldsymbol{B}_m^{-1}\boldsymbol{V}^\top$, from which we have $\boldsymbol{C}_m = \boldsymbol{V}\boldsymbol{D}_m\boldsymbol{V}^\top$, where

$$\boldsymbol{D}_m = \lambda^{-1}(\boldsymbol{B}_m^{-1} - \boldsymbol{I}_n). \tag{49}$$

Let

$$d_i^{(m)} = \frac{1}{\lambda}\left(\frac{1}{b_i^{(m)}} - 1\right), \quad i = 1, \ldots, n. \tag{50}$$

Then, $\boldsymbol{D}_m = \mathrm{diag}(d_1^{(m)}, \ldots, d_n^{(m)})$ and from the inequalities given in (13), it follows that

$$0 = d_1^{(m)} = d_2^{(m)} < d_3^{(m)} \leq \cdots \leq d_n^{(m)}. \tag{51}$$

10. Since $\boldsymbol{A}_m\boldsymbol{\Pi} = \boldsymbol{\Pi}$, it follows that

$$\boldsymbol{A}_m^{-1}\boldsymbol{\Pi} = \boldsymbol{\Pi},$$

which leads to

$$\boldsymbol{C}_m\boldsymbol{\Pi} = \lambda^{-1}(\boldsymbol{A}_m^{-1} - \boldsymbol{I}_n)\boldsymbol{\Pi} = \lambda^{-1}(\boldsymbol{\Pi} - \boldsymbol{\Pi}) = \boldsymbol{0}. \tag{52}$$

Thus, both $\boldsymbol{\iota}$ and $\boldsymbol{\tau}$ belong to the null space of $\boldsymbol{C}_m$. In addition, given that $0 < b_i^{(m)} < 1$ for $i = 3, \ldots, n$, from (49), $\boldsymbol{C}_m$ is a nonnegative definite matrix whose rank is $n - 2$.

11. Let $\boldsymbol{\theta} = \boldsymbol{V}^\top \boldsymbol{x}$. Then, since $\boldsymbol{x} = \boldsymbol{V}\boldsymbol{\theta}$, $f^{(m)}(\boldsymbol{x})$ can be represented by $\boldsymbol{\theta}$ as

$$\begin{aligned} f^{(m)}(\boldsymbol{x}) &= \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \boldsymbol{x}^\top \boldsymbol{C}_m \boldsymbol{x} = \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \boldsymbol{x}^\top \boldsymbol{V}\boldsymbol{D}_m\boldsymbol{V}^\top \boldsymbol{x} \\ &= \|\boldsymbol{y} - \boldsymbol{V}\boldsymbol{\theta}\|^2 + \lambda \boldsymbol{\theta}^\top \boldsymbol{D}_m \boldsymbol{\theta}. \end{aligned}$$

12. Let us consider the following generalized ridge regression:

$$\min_{\boldsymbol{\theta}} g^{(m)}(\boldsymbol{\theta}) = \|\boldsymbol{y} - \boldsymbol{V}\boldsymbol{\theta}\|^2 + \lambda \boldsymbol{\theta}^\top \boldsymbol{D}_m \boldsymbol{\theta}. \tag{53}$$

From (51), the first two entries of $\boldsymbol{\theta}$ are not penalized. Let

$$\widehat{\boldsymbol{\theta}}^{(m)} = (\boldsymbol{I}_n + \lambda \boldsymbol{D}_m)^{-1} \boldsymbol{V}^\top \boldsymbol{y}. \tag{54}$$

Then, it follows that $g^{(m)}(\boldsymbol{\theta}) > g^{(m)}(\widehat{\boldsymbol{\theta}}^{(m)})$ for all $\boldsymbol{\theta} \in \mathbb{R}^n \backslash \{\widehat{\boldsymbol{\theta}}^{(m)}\}$. It also follows that

$$\begin{aligned} \boldsymbol{V}\widehat{\boldsymbol{\theta}}^{(m)} &= \boldsymbol{V}(\boldsymbol{I}_n + \lambda \boldsymbol{D}_m)^{-1} \boldsymbol{V}^\top \boldsymbol{y} = (\boldsymbol{I}_n + \lambda \boldsymbol{V}\boldsymbol{D}_m\boldsymbol{V}^\top)^{-1} \boldsymbol{y} \\ &= (\boldsymbol{I}_n + \lambda \boldsymbol{C}_m)^{-1} \boldsymbol{y} = \boldsymbol{A}_m \boldsymbol{y} = \widehat{\boldsymbol{x}}^{(m)}. \end{aligned} \tag{55}$$

Let us summarize the above results.

Proposition 1. *(i) The mean of the entries of $\widehat{\boldsymbol{x}}^{(m)}$ is equal to that of $\boldsymbol{y}$. In addition, the bHP filter trend residuals, $\boldsymbol{y} - \widehat{\boldsymbol{x}}^{(m)}$, sum to zero. (ii) Each row of the smoother matrix $\boldsymbol{A}_m$ sums to unity. (iii) If $\boldsymbol{\xi} \in \mathbb{S}(\boldsymbol{\Pi})$, then $\boldsymbol{A}_m\boldsymbol{\xi} = \boldsymbol{\xi}$. In addition, it follows that $\widehat{\boldsymbol{x}}^{(m)} = \boldsymbol{P}_{\boldsymbol{\Pi}}\boldsymbol{y} + \boldsymbol{A}_m(\boldsymbol{y} - \boldsymbol{P}_{\boldsymbol{\Pi}}\boldsymbol{y})$. (iv) When n and m are fixed, as $\lambda \to \infty$, $\widehat{\boldsymbol{x}}^{(m)} \to \boldsymbol{P}_{\boldsymbol{\Pi}}\boldsymbol{y}$. (v) When n and m are fixed, as $\lambda \to 0$, $\widehat{\boldsymbol{x}}^{(m)} \to \boldsymbol{y}$. (vi) When n and λ are fixed, as $m \to \infty$, $\widehat{\boldsymbol{x}}^{(m)} \to \boldsymbol{y}$. (vii) When n, m, and λ are fixed, as $k \to \infty$, $\boldsymbol{A}_m^k\boldsymbol{y} \to \boldsymbol{P}_{\boldsymbol{\Pi}}\boldsymbol{y}$. (viii) $\widehat{\boldsymbol{x}}^{(m)}$ is the solution of the penalized least squares problem given by (48). The penalty matrix of the bHP filter can be represented by $\boldsymbol{C}_m$, which is an $n \times n$ nonnegative definite matrix whose null space is identical to $\mathbb{S}(\boldsymbol{\Pi})$. (ix) (53) is a generalized ridge regression representation of the bHP filter.*

Remark 1. Regarding Proposition 1, we make five remarks:

1. Some of the results in Proposition 1 are generalizations of those in [12, Proposition 2.2], which documents several properties of the HP filter. For example, Proposition 1(i) is a generalization of [12, Proposition 2.2(iii)(a)].
2. Given that $\boldsymbol{A}_m$ is symmetric, from Proposition 1(ii), it follows that $\boldsymbol{\iota}^\top \boldsymbol{A}_m = \boldsymbol{\iota}^\top$. Then, Proposition 1(i) can also be derived from it as

$$\boldsymbol{\iota}^\top \widehat{\boldsymbol{x}}^{(m)} = \boldsymbol{\iota}^\top \boldsymbol{A}_m \boldsymbol{y} = \boldsymbol{\iota}^\top \boldsymbol{y}.$$

3. $\widehat{\boldsymbol{x}}^{(m)} = \boldsymbol{P}_{\boldsymbol{\Pi}}\boldsymbol{y} + \boldsymbol{A}_m(\boldsymbol{y} - \boldsymbol{P}_{\boldsymbol{\Pi}}\boldsymbol{y})$ in Proposition 1(iii) is a generalization of the result in [13, p. 342] and [14, Eq. (3)]. Given that $\boldsymbol{A}_m$ is a low-pass filter, it indicates that $\widehat{\boldsymbol{x}}^{(m)}$ is the sum of the linear time trend estimated by OLS, $\boldsymbol{P}_{\boldsymbol{\Pi}}\boldsymbol{y}$, and low-frequency components in the linear trend residuals, $\boldsymbol{y} - \boldsymbol{P}_{\boldsymbol{\Pi}}\boldsymbol{y}$.

Table 1. List of the relationships between the main matrices

	HP filter	bHP filter
Smoother matrix	$\boldsymbol{A}$	$\boldsymbol{A}_m$
Spectral decomposition of smoother matrix	$\boldsymbol{V}\boldsymbol{B}\boldsymbol{V}^\top$	$\boldsymbol{V}\boldsymbol{B}_m\boldsymbol{V}^\top$
Penalty matrix	$\boldsymbol{C}$	$\boldsymbol{C}_m$
Spectral decomposition of penalty matrix	$\boldsymbol{V}\boldsymbol{D}\boldsymbol{V}^\top$	$\boldsymbol{V}\boldsymbol{D}_m\boldsymbol{V}^\top$

4. Since $\widehat{\boldsymbol{x}}^{(m)} = \boldsymbol{A}_m\boldsymbol{y}$, for example, $\boldsymbol{A}_m^2\boldsymbol{y} = \boldsymbol{A}_m\widehat{\boldsymbol{x}}^{(m)}$. Thus, $\boldsymbol{A}_m^k\boldsymbol{y}$ in Proposition 1(vii) represents the result of k times repeated bHP filtering. In addition, from Proposition 1(vii), it follows that

$$\boldsymbol{\Delta}_2(\boldsymbol{A}_m^k\boldsymbol{y}) \to \boldsymbol{0}, \quad (n, m, \lambda : \text{fixed},\ k \to \infty), \tag{56}$$

which is a generalization of the result in [2, p. 961].

5. Proposition 1(viii) and (ix) are restatements of the results in [4]. Given that $\boldsymbol{A}_1 = \boldsymbol{A} = (\boldsymbol{I}_n + \lambda\boldsymbol{C})^{-1}$, by its definition, it follows that

$$\boldsymbol{C}_1 = \lambda^{-1}(\boldsymbol{A}_1^{-1} - \boldsymbol{I}_n) = \lambda^{-1}(\boldsymbol{I}_n + \lambda\boldsymbol{C} - \boldsymbol{I}_n) = \boldsymbol{C}. \tag{57}$$

Therefore, if $m = 1$, then (48) becomes

$$\min_{\boldsymbol{x}} f^{(1)}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda\boldsymbol{x}^\top\boldsymbol{C}\boldsymbol{x}, \tag{58}$$

which is the HP filter. From (15), (19), and (50), it also follows that

$$d_i^{(1)} = \frac{1}{\lambda}\left(\frac{1}{b_i^{(1)}} - 1\right) = \frac{1}{\lambda}\left(\frac{1}{b_i} - 1\right) = d_i, \quad i = 1, \ldots, n. \tag{59}$$

from which we obtain

$$\boldsymbol{D}_1 = \boldsymbol{D}. \tag{60}$$

This is consistent with (57).

5 Concluding Remarks

In this paper, we have derived several properties of the bHP filter. For this purpose, we used the spectral decomposition of the penalty matrix of the HP filter. The results of this paper are summarized in Proposition 1. See also Table 1, which lists the relationships between the main matrices such as $\boldsymbol{A}_m$ and $\boldsymbol{B}_m$.

Finally, we make a remark. The HP filter is a kind of the WH(p) graduation defined by

$$\min_{x_1,\ldots,x_n} f_p(x_1, \ldots, x_n) = \sum_{i=1}^{n}(y_i - x_i)^2 + \lambda_p \sum_{i=p+1}^{n}(\Delta^p x_i)^2, \tag{61}$$

where $\lambda_p \in (0, \infty)$ is a smoothing parameter and $\Delta x_i = x_i - x_{i-1}$. (Note that $\Delta^2 x_i = \Delta(\Delta x_i) = x_i - 2x_{i-1} + x_{i-2}$.) Then, we can consider the boosted version of the WH(p) graduation and establish its properties using the approach adopted in this paper. We are investigating this issue and will report our findings in the future.

Acknowledgments. The author would like to thank three anonymous reviewers for their valuable comments and suggestions. He also thanks Keith Knight for permission to reference his unpublished manuscript. This study was funded by the Japan Society for the Promotion of Science (KAKENHI Grant Number 23K013770A).

Appendix: Some Proofs

Proof of (12)
$\boldsymbol{C}$ in (5) is the following $n \times n$ symmetric pentadiagonal matrix.

$$\boldsymbol{C} = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & \cdots & 0 \\ -2 & 5 & -4 & 1 & \ddots & & \vdots \\ 1 & -4 & 6 & \ddots & \ddots & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & \ddots & 1 & 0 \\ \vdots & \ddots & \ddots & \ddots & 6 & -4 & 1 \\ \vdots & & \ddots & 1 & -4 & 5 & -2 \\ 0 & \cdots & \cdots & 0 & 1 & -2 & 1 \end{bmatrix}.$$

From the Gershgorin circle theorem, its eigenvalues, d_i for $i = 1, \ldots, n$, satisfy

$$d_i \in L_1 \cup L_2 \cup L_3, \quad i = 1, \ldots, n,$$

where $L_1 = \{\eta : |\eta - 1| \leq 3\}$, $L_2 = \{\eta : |\eta - 5| \leq 7\}$, and $L_3 = \{\eta : |\eta - 6| \leq 10\}$. Therefore, $d_n \leq 16$ follows.
Another Proof of (12)

Kim et al. [13, p. 342] show $\|\boldsymbol{\Delta}_2 \boldsymbol{x}\| \leq 4\|\boldsymbol{x}\|$, from which we obtain

$$\frac{\|\boldsymbol{\Delta}_2 \boldsymbol{x}\|^2}{\|\boldsymbol{x}\|^2} \leq 16.$$

If $\boldsymbol{x} = \boldsymbol{v}_n$, then the left-hand side is equal to d_n.
Proof of (26)

It follows that

$$b_1^{(m)} z_1 \boldsymbol{v}_1 + b_2^{(m)} z_2 \boldsymbol{v}_2 = \boldsymbol{V}_1 \boldsymbol{V}_1^\top \boldsymbol{y} = \boldsymbol{V}_1 (\boldsymbol{V}_1^\top \boldsymbol{V}_1)^{-1} \boldsymbol{V}_1^\top \boldsymbol{y} = \boldsymbol{P}_{V_1} \boldsymbol{y},$$

which is the orthogonal projection of $\boldsymbol{y}$ onto $\mathbb{S}(\boldsymbol{V}_1)$. On the other hand, $\boldsymbol{P}_{\Pi} \boldsymbol{y}$ is the orthogonal projection of $\boldsymbol{y}$ onto $\mathbb{S}(\boldsymbol{\Pi})$. Since $\mathbb{S}(\boldsymbol{V}_1)$ and $\mathbb{S}(\boldsymbol{\Pi})$ are identical, (26) holds.

References

1. Hodrick, R.J., Prescott, E.C.: Postwar U.S. business cycles: an empirical investigation. J. Money Credit Banking **29**(1), 1–16 (1997)
2. Weinert, H.L.: Efficient computation for Whittaker-Henderson smoothing. Comput. Stat. Data Anal. **52**(2), 959–974 (2007)
3. Phillips, P.C.B., Shi, Z.: Boosting: why you can use the HP filter. Int. Econ. Rev. **62**(2), 521–570 (2021)
4. Knight, K.: The boosted Hodrick–Prescott filter, penalized least squares, and Bernstein polynomials. Department of Statistical Sciences, University of Toronto (2021). https://www.utstat.utoronto.ca/keith/papers/hp-pls.pdf
5. Biswas, E, Sabzikar, F., Phillips, P.C.B.: Boosting the HP filter for trending time series with long range dependence. Cowles Foundation Discussion Papers No. 2347, Yale University (2022). https://cowles.yale.edu/sites/default/files/2022-12/d2347.pdf
6. Hall, V., Thomson, P.: A boosted HP filter for business cycle analysis: evidence from New Zealand's small open economy. SEF Working Paper 1/2022, Victoria University of Wellington (2022). https://researcharchive.vuw.ac.nz/xmlui/bitstream/handle/10063/9473/Working%20Paper.pdf?sequence=6
7. Mei, Z., Phillips, P.C.B., Shi, Z.: The boosted HP filter is more general than you might think. arXiv:2209.09810 (2022). https://doi.org/10.48550/arXiv.2209.09810
8. Tomal, M.: Testing for overall and cluster convergence of housing rents using robust methodology: evidence from Polish provincial capitals. Empirical Econ. **62**, 2023–2055 (2022)
9. Widiantoro, D.M.: Countercyclical capital buffer: building the resilience or taming the rapid financial cycle? MPRA Paper No. 113507 (2022). https://mpra.ub.uni-muenchen.de/113507/
10. Lu, Y., Pagan, A.: To boost or not to boost? That is the question. CAMA Working Paper 12/2023, Centre for Applied Macroeconomic Analysis, Australian National University (2023). https://crawford.anu.edu.au/sites/default/files/publication/cama_crawford_anu_edu_au/2023-02/12_2023_lu_pagan_0.pdf
11. Danthine, J.-P., Girardin, M.: Business cycles in Switzerland: a comparative study. Eur. Econ. Rev. **33**(1), 31–50 (1989)
12. Yamada, H.: A smoothing method that looks like the Hodrick-Prescott filter. Economet. Theor. **36**(5), 961–981 (2020)
13. Kim, S., Koh, K., Boyd, S., Gorinevsky, D.: ℓ_1 trend filtering. SIAM Rev. **51**(2), 339–360 (2009)
14. Yamada, H.: Why does the trend extracted by the Hodrick-Prescott filtering seem to be more plausible than the linear trend? Appl. Econ. Lett. **25**(2), 102–105 (2018)
15. Abadir, K.M., Magnus, J.R.: Matrix Algebra. Cambridge University Press, Cambridge (2005)