Contemporary Mathematics
https://ojs.wiserpub.com/index.php/CM/

Research Article

Boosted Whittaker-Henderson Graduation of Order 1: A Graph Spectral Filter Using Discrete Cosine Transform

Ruoyi Bao, **Hiroshi Yamada***

Graduate School of Humanities and Social Sciences, Hiroshima University, 1-2-1 Kagamiyama, Higashihiroshima 739-8525, Japan
E-mail: yamada@hiroshima-u.ac.jp

Received: 2 February 2024; **Revised:** 28 March 2024; **Accepted:** 7 April 2024

Abstract: The Whittaker-Henderson (WH) graduation of order 1 is a smoothing/filtering method for equally spaced one-dimensional data. Inspired by Phillips and Shi, this paper introduces the boosted version of the WH graduation of order 1. We show that it is a graph spectral filter using the discrete cosine transform, and then provide a simple formula for the (i, j) entry of its smoother matrix. We also show that it is a linear smoother such that the filter weights sum to unity and the smoother matrix is bisymmetric, i.e., symmetric and centrosymmetric. GNU Octave user-defined functions based on the obtained results are also provided.

***Keywords*:** Whittaker-Henderson graduation, boosted Hodrick-Prescott filter, graph Laplacian, graph spectral filter, discrete cosine transform

MSC: 62G05

1. Introduction

The Whittaker-Henderson (WH) graduation of order 1, or WH(1) graduation for short, is a smoothing/filtering method for equally spaced one-dimensional data. The filter has a long history, and its origin can be traced back to Bohlmann [1] at least. For the history of the WH graduation, see Weinert [2]. The WH graduation of order 2 is called the Hodrick-Prescott (HP) [3] filter in the field of econometrics and it is a prominent method for trend estimation of macroeconomic time series data. Roughly speaking, the HP filter is suitable for smoothing data with a linear trend, while the WH(1) graduation is suitable for smoothing data without a trend.

Recently, Phillips and Shi [4] made a major improvement to the HP filter. By adapting the L_2-boosting procedure in machine learning to the HP filter, they developed a boosted HP (bHP) filter. The new filter is designed to recover the trend elements retained in the trend residuals. Several studies concerning the bHP filter have since emerged, including Knight [5], Biswas et al. [6], Hall and Thomson [7], Mei et al. [8], Tomal [9], Widiantoro [10], and Lu and Pagan [11].

Inspired by Phillips and Shi [4], as the bHP filter was developed from the HP filter, in this paper, we introduce a boosted version of the WH graduation of order 1. We shall call it boosted WH (1) graduation or bWH (1) graduation for short. Roughly speaking, as before, the bHP filter is suitable for smoothing data with a linear trend, while the bWH(1) graduation is suitable for smoothing data without a trend. After defining it, we show that it is a graph spectral filter

DOI: https://doi.org/10.37256/cm.6220254409

Shuman, et al. [12] using the discrete cosine transform (DCT) Ahmed et al. [13]. Then, we provide a simple formula for the $(i,\ j)$ entry of its smoother matrix. We also show that it is a linear smoother such that the filter weights sum to unity and the smoother matrix is bisymmetric, i.e., symmetric and centrosymmetric. GNU Octave user-defined functions based on the obtained results are also provided.

Here, we mention related studies. Yamada and Jahra [14] studied the WH(1) graduation and derived a formula for the $(i,\ j)$ entry of its smoother matrix. Subsequently, Yamada [15, Proposition 4.10 (iv)] provided a simpler formula for it. In addition, Yamada [16, Theorem 2.2] showed that the smoother matrix of the WH(1) graduation is bisymmetric. This paper includes generalized results from Yamada [15, 16].

The organization of the paper is as follows. In Section 2, we introduce the bWH(1) graduation and provide some remarks about it. In Section 3, we present the main results. In Section 4, we give some reasons why the bWH(1) graduation is required. Section 5 concludes the paper. In the Appendix, we provide proofs and GNU Octave user-defined functions.

Table 1. List of acronyms

WH graduation	Whittaker-Henderson graduation
WH(1) graduation	Whittaker-Henderson graduation of order 1
HP filter	Hodrick-Prescott filter
bHP filter	boosted Hodrick-Prescott filter
bWH(1) graduation	boosted Whittaker-Henderson graduation of order 1
DCT	discrete cosine transform
OLS	ordinary least squares

2. The boosted WH(1) graduation

The WH(1) graduation is defined by the following minimization problem:

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

where $\lambda \in [0,\ \infty)$ is a smoothing parameter that controls fidelity, $\sum_{i=1}^{n}(y_i - x_i)^2$, and smoothness, $\sum_{i=2}^{n}(x_i - x_{i-1})^2$. Let $\boldsymbol{y} = [y_1,\ \dots,\ y_n]^\top$, $\boldsymbol{x} = [x_1,\ \dots,\ x_n]^\top$, $\boldsymbol{I}_n$ be the identity matrix of order n, and $\boldsymbol{D}$ be the $(n-1)\times n$ matrix such that $\boldsymbol{D}\boldsymbol{x} = [x_2 - x_1,\ \dots,\ x_n - x_{n-1}]^\top$. In addition, for a vector $\boldsymbol{\xi} = [\xi_1,\ \dots,\ \xi_l]^\top$, let $\|\boldsymbol{\xi}\|^2 = \sum_{i=1}^{l}\xi_i^2$. Then, (1) can be represented in matrix form as

$$\min_{\boldsymbol{x}} \ f(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda\|\boldsymbol{D}\boldsymbol{x}\|^2 = \boldsymbol{x}^\top(\boldsymbol{I}_n + \lambda\boldsymbol{D}^\top\boldsymbol{D})\boldsymbol{x} - 2\boldsymbol{y}^\top\boldsymbol{x} + \boldsymbol{y}^\top\boldsymbol{y}. \tag{2}$$

Since $f(\boldsymbol{x})$ is a quadratic function of $\boldsymbol{x}$ whose Hessian matrix, $2(\boldsymbol{I}_n + \lambda\boldsymbol{D}^\top\boldsymbol{D})$, is positive definite, there exists $\widetilde{\boldsymbol{x}}$ such that $f(\boldsymbol{x}) > f(\widetilde{\boldsymbol{x}})$ for all $\boldsymbol{x} \in \mathbb{R}^n \setminus \{\widetilde{\boldsymbol{x}}\}$. Explicitly, it is given by $\widetilde{\boldsymbol{x}} = \boldsymbol{S}\boldsymbol{y}$, where $\boldsymbol{S} = (\boldsymbol{I}_n + \lambda\boldsymbol{D}^\top\boldsymbol{D})^{-1}$, which is the smoother matrix of the WH(1) graduation.

Let us define the bWH(1) graduation. It is defined by

$$\widehat{\boldsymbol{x}}^{(m)} = \boldsymbol{S}^{(m)}\boldsymbol{y}. \tag{3}$$

Here, $\boldsymbol{S}^{(m)} = \boldsymbol{I}_n - (\boldsymbol{I}_n - \boldsymbol{S})^m$, where m is a positive integer. Note that since $\widehat{\boldsymbol{x}}^{(1)} = \boldsymbol{S}^{(1)}\boldsymbol{y} = \{\boldsymbol{I}_n - (\boldsymbol{I}_n - \boldsymbol{S})\}\boldsymbol{y} = \boldsymbol{S}\boldsymbol{y} = \widetilde{\boldsymbol{x}}$, it is a generalization of the WH(1) graduation.

3. The main results

In this section, we show that the bWH(1) graduation is a type of graph spectral filter such that its graph Fourier transform is the DCT. Then, we provide an explicit formula for the $(i,\ j)$ entry of $\boldsymbol{S}^{(m)}$. We also show that the bWH(1) graduation is a linear smoother whose filter weights sum to unity and $\boldsymbol{S}^{(m)}$ is bisymmetric.

Consider the undirected graph $G = (V,\ E)$ such that $V = \{v_1,\ \dots, v_n\}$ and $E = \{\{v_1,\ v_2\},\ \dots,\ \{v_{n-1},\ v_n\}\}$, which is a path graph, and denote its graph Laplacian by $\boldsymbol{L}$. Then, $\boldsymbol{L}$ is an $n \times n$ tridiagonal symmetric matrix (e.g., Nakatsukasa et al. [17], Strang and MacNamara [18, Eq. (9.1)]). Since $\boldsymbol{D}^\top$ is an incidence matrix of $G = (V,\ E)$, it follows that $\boldsymbol{L} = \boldsymbol{D}^\top\boldsymbol{D}$. Then, the smoother matrix of the WH(1) graduation, $\boldsymbol{S}$, can be represented in term of $\boldsymbol{L}$ as

$$\boldsymbol{S} = (\boldsymbol{I}_n + \lambda \boldsymbol{L})^{-1}. \tag{4}$$

Let $\boldsymbol{Q} = \operatorname{diag}(q_1,\ \dots,\ q_n)$ and $\boldsymbol{P} = [\boldsymbol{p}_1,\ \dots,\ \boldsymbol{p}_n]$, where $(q_i,\ \boldsymbol{p}_i)$ for $i = 1,\ \dots,\ n$ are the eigenpairs of a real symmetric matrix $\boldsymbol{L}$ such that $q_1,\ \dots,\ q_n$ are in ascending order and $\boldsymbol{p}_1,\ \dots,\ \boldsymbol{p}_n$ are orthonormal eigenvectors. Then, from, e.g., von Neumann [19], q_i for $i = 1,\ \dots,\ n$ are explicitly represented as

$$q_i = 2 - 2\cos\left\{\frac{(i-1)\pi}{n}\right\} = 4\sin^2\left\{\frac{(i-1)\pi}{2n}\right\}, \quad i = 1,\ \dots,\ n. \tag{5}$$

In addition, we can let $\boldsymbol{p}_1 = \frac{1}{\sqrt{n}}\boldsymbol{\iota}$, where $\boldsymbol{\iota}$ is the n-dimensional vector of ones, and

$$\boldsymbol{p}_i = \sqrt{\frac{2}{n}}\left[\cos\left\{(i-1)\theta_1\right\},\ \cos\left\{(i-1)\theta_2\right\},\ \dots,\ \cos\left\{(i-1)\theta_n\right\}\right]^\top, \quad i = 2,\ \dots,\ n,$$

where $\theta_h = \frac{\left(h - \frac{1}{2}\right)\pi}{n}$ for $h = 1,\ \dots,\ n$.

Remark 1 We give three remarks on the eigenpairs.

(i) $\boldsymbol{P}\boldsymbol{Q}\boldsymbol{P}^\top$ is a spectral decomposition of $\boldsymbol{L}$.

(ii) $\boldsymbol{P}^\top\boldsymbol{y}$ is the DCT of $\boldsymbol{y}$ (e.g., Strang [20], Yamada, [21]).

(iii) It follows that $0 = q_1 < q_2 < \cdots < q_n < 4$. Accordingly, q_2 is the Fiedler eigenvalue and the rank of $\boldsymbol{L}$ is $n-1$, which is consistent with that the number of connected components of $G = (V,\ E)$ is unity (see, e.g., Bapat [22], Gallier [23]).

Let $\boldsymbol{\Phi} = \operatorname{diag}(\phi_1,\ \dots,\ \phi_n)$, where $\phi_i = 1 - \left(\frac{\lambda q_i}{1 + \lambda q_i}\right)^m$ for $i = 1,\ \dots,\ n$. Here, given that $\lambda > 0$, $0 = q_1 < q_2 < \cdots < q_n$, and m is a positive integer, it follows that

$$1 = \phi_1 > \phi_2 > \cdots > \phi_n > 0. \tag{6}$$

For a proof of (6), see Appendix A.1.

Given that $\boldsymbol{L} = \boldsymbol{PQP}^\top$ and $\boldsymbol{P}$ is an orthogonal matrix, it follows that $\boldsymbol{S} = (\boldsymbol{I}_n + \lambda \boldsymbol{L})^{-1} = (\boldsymbol{PP}^\top + \lambda \boldsymbol{PQP}^\top)^{-1} = \boldsymbol{P}(\boldsymbol{I}_n + \lambda \boldsymbol{Q})^{-1}\boldsymbol{P}^\top$. Accordingly, the smoother matrix of the bWH(1) graduation $\boldsymbol{S}^{(m)}$ can be represented in term of $\boldsymbol{P}$ and $\boldsymbol{\Phi}$ as follows.

$$\begin{aligned}\boldsymbol{S}^{(m)} &= \boldsymbol{I}_n - (\boldsymbol{I}_n - \boldsymbol{S})^m = \boldsymbol{PP}^\top - \left\{\boldsymbol{PP}^\top - \boldsymbol{P}(\boldsymbol{I}_n + \lambda \boldsymbol{Q})^{-1}\boldsymbol{P}^\top\right\}^m \\ &= \boldsymbol{PP}^\top - \boldsymbol{P}\left\{\boldsymbol{I}_n - (\boldsymbol{I}_n + \lambda \boldsymbol{Q})^{-1}\right\}^m \boldsymbol{P}^\top \\ &= \boldsymbol{P}\left[\boldsymbol{I}_n - \left\{\boldsymbol{I}_n - (\boldsymbol{I}_n + \lambda \boldsymbol{Q})^{-1}\right\}^m\right]\boldsymbol{P}^\top = \boldsymbol{P\Phi P}^\top.\end{aligned} \tag{7}$$

Here, the last equality in (7) follows because $\boldsymbol{\Phi} = \boldsymbol{I}_n - \left\{\boldsymbol{I}_n - (\boldsymbol{I}_n + \lambda \boldsymbol{Q})^{-1}\right\}^m$. Note that $\boldsymbol{P\Phi P}^\top$ in (7) is a spectral decomposition of $\boldsymbol{S}^{(m)}$.

Consequently, we obtain the following result.

Proposition 1 The boosted WH(1) graduation is a graph spectral filter such that it is given by $\widehat{\boldsymbol{x}}^{(m)} = \boldsymbol{P\Phi P}^\top \boldsymbol{y}$.

Proof. It immediately follows from (7) . □

Remark 2 Concerning Proposition 1, we give three remarks.

(i) The calculation of $\boldsymbol{P\Phi P}^\top \boldsymbol{y}$ can be decomposed as the following three steps:

Steps 1 $\boldsymbol{P}^\top \boldsymbol{y}$: DCT of $\boldsymbol{y}$.

Steps 2 $\boldsymbol{\Phi}(\boldsymbol{P}^\top \boldsymbol{y})$: low-pass filtering of $(\boldsymbol{P}^\top \boldsymbol{y})$.

Steps 3 $\boldsymbol{P}(\boldsymbol{\Phi P}^\top \boldsymbol{y})$: inverse DCT of $(\boldsymbol{\Phi P}^\top \boldsymbol{y})$.

Note that we can also see the same structure in the WH(1) graduation and in the filter considered by Garcia [24]. Here, Garcia's [24] filter is given by $\boldsymbol{x}^\dagger = (\boldsymbol{I}_n + \lambda \boldsymbol{L}^2)^{-1}\boldsymbol{y}$.

(ii) Let $\eta(\boldsymbol{y})$ denote the von Neumann's [19] ratio of $\boldsymbol{y}$:

$$\eta(\boldsymbol{y}) = \frac{\sum_{i=2}^n (y_i - y_{i-1})^2/(n-1)}{\sum_{i=1}^n (y_i - \bar{y})^2/n} = \frac{n}{n-1}\frac{\boldsymbol{y}^\top \boldsymbol{L}\boldsymbol{y}}{\boldsymbol{y}^\top \boldsymbol{M}\boldsymbol{y}}, \tag{8}$$

where $\bar{y} = \frac{1}{n}\sum_{i=1}^n y_i$ and $\boldsymbol{M} = \boldsymbol{I}_n - \boldsymbol{\iota}(\boldsymbol{\iota}^\top\boldsymbol{\iota})^{-1}\boldsymbol{\iota}^\top$. Then, given that (a) $\boldsymbol{p}_i^\top \boldsymbol{L}\boldsymbol{p}_i = q_i$ and $\boldsymbol{p}_i^\top \boldsymbol{M}\boldsymbol{p}_i = 1$ for $i = 2, \ldots, n$ and (b) $0 < q_2 < \cdots < q_n$, it follows that

$$0 < \eta(\boldsymbol{p}_2) < \cdots < \eta(\boldsymbol{p}_n), \tag{9}$$

which implies that $\boldsymbol{p}_i$ is smoother than $\boldsymbol{p}_{i+1}$ for $i = 2, \ldots, n-1$. Incidentally, the von Neumann's ratio $\eta(\boldsymbol{y})$ in (8) can be regarded as a type of Geary's c, which is a prominent spatial autocorrelation measure. Accordingly, denoting the Geary's c of $\boldsymbol{y}$ by $c(\boldsymbol{y})$, it follows that $0 < c(\boldsymbol{p}_2) < \cdots < c(\boldsymbol{p}_n)$. For details on the Geary's c, see, e.g., Yamada [25], [26] and the references therein.

(iii) A GNU Octave user-defined function to calculate $\widehat{\boldsymbol{x}}^{(m)}$ based on Proposition 1, `calc_xhat_dct`, is provided in Appendix B. 1. It utilizes `dct` and `idct` functions.

Given that $\phi_1 = 1$ and $\boldsymbol{p}_1\boldsymbol{p}_1^\top = \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top$, it follows that

$$\boldsymbol{S}^{(m)} = \boldsymbol{P}\boldsymbol{\Phi}\boldsymbol{P}^\top = \sum_{k=1}^{n} \phi_k \boldsymbol{p}_k \boldsymbol{p}_k^\top = \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top + \sum_{k=2}^{n} \phi_k \boldsymbol{p}_k \boldsymbol{p}_k^\top. \tag{10}$$

Consequently, we obtain the following result.

Proposition 2 Denote the $(i,\ j)$ entry of $\boldsymbol{S}^{(m)}$ by $s_{i,\ j}^{(m)}$. Then, it follows that

$$s_{i,\ j}^{(m)} = \frac{1}{n} + \frac{2}{n}\sum_{k=2}^{n} \phi_k \cos\{(k-1)\theta_i\}\cos\{(k-1)\theta_j\}, \quad i,\ j = 1,\ \ldots,\ n. \tag{11}$$

Proof. It immediately follows from (10). □

From Proposition 2, we have the following results.

Corollary 1 The row/column sum of $\boldsymbol{S}^{(m)}$ equals unity.

Proof. Given that $\boldsymbol{p}_1^\top \boldsymbol{p}_k = 0$ for $k = 2,\ \ldots,\ n$, it follows that $\frac{\sqrt{2}}{n}\sum_{j=1}^{n}\cos\{(k-1)\theta_j\} = 0$. Then, from Proposition 2, we obtain

$$\sum_{j=1}^{n} s_{i,\ j}^{(m)} = \sum_{j=1}^{n}\frac{1}{n} + \frac{2}{n}\sum_{k=2}^{n}\phi_k\cos\{(k-1)\theta_i\}\sum_{j=1}^{n}\cos\{(k-1)\theta_j\} = 1.$$

Given that $\boldsymbol{S}^{(m)}$ is symmetric, $\sum_{j=1}^{n} s_{i,\ j}^{(m)} = 1$ implies $\sum_{i=1}^{n} s_{i,\ j}^{(m)} = 1$.

Corollary 2 $\boldsymbol{S}^{(m)}$ is bisymmetric.

Proof. From (11), it immediately follows that $s_{i,\ j}^{(m)} = s_{j,\ i}^{(m)}$ for $i,\ j = 1,\ \ldots,\ n$. Thus, $\boldsymbol{S}^{(m)}$ is symmetric. Next, we show that $\boldsymbol{S}^{(m)}$ is centrosymmetric. Given that $\theta_h = \frac{\left(h - \frac{1}{2}\right)\pi}{n}$, it follows that

$$\cos\{(k-1)\theta_{n-i+1}\} = \cos\left\{(k-1)\frac{\left(n-i+1-\frac{1}{2}\right)\pi}{n}\right\} = \cos\{(k-1)\pi - (k-1)\theta_i\}$$

$$= \cos\{(k-1)\pi\}\cos\{(k-1)\theta_i\} + \sin\{(k-1)\pi\}\sin\{(k-1)\theta_i\}$$

$$= \cos\{(k-1)\pi\}\cos\{(k-1)\theta_i\}.$$

Accordingly, given that $\cos^2\{(k-1)\pi\} = 1$, we have

$$\cos\{(k-1)\theta_{n-i+1}\}\cos\{(k-1)\theta_{n-j+1}\} = \cos^2\{(k-1)\pi\}\cos\{(k-1)\theta_i\}\cos\{(k-1)\theta_j\}$$

$$= \cos\{(k-1)\theta_i\}\cos\{(k-1)\theta_j\}, \quad i,\ j = 1,\ \ldots,\ n.$$

Thus, it follows that

$$s^{(m)}_{n-i+1,\ n-j+1} = \frac{1}{n} + \frac{2}{n}\sum_{k=2}^{n} \phi_k \cos\{(k-1)\theta_{n-i+1}\}\cos\{(k-1)\theta_{n-j+1}\}$$

$$= \frac{1}{n} + \frac{2}{n}\sum_{k=2}^{n} \phi_k \cos\{(k-1)\theta_i\}\cos\{(k-1)\theta_j\} = s^{(m)}_{i,\ j} \quad i,\ j = 1,\ \ldots,\ n,$$

which proves that $\boldsymbol{S}^{(m)}$ is centrosymmetric. Therefore, $\boldsymbol{S}^{(m)}$ is bisymmetric.

Remark 3 Concerning Proposition 2 and Corollaries 1-2, we give five remarks.

(i) Proposition 2 (resp. Corollary 2) can be regarded as a generalization of Yamada [15, Proposition 4.10 (iv)] (resp. a part of Yamada [16, Theorem 2.2]).

(ii) Denote the i-th entry of $\widehat{\boldsymbol{x}}^{(m)}$ by $\widehat{x}^{(m)}_i$.

Then, from Corollary 1, it follows that

$$\widehat{x}^{(m)}_i = s^{(m)}_{i,\ 1} y_1 + \cdots + s^{(m)}_{i,\ n} y_n, \tag{12}$$

where $\sum_{j=1}^{n} s^{(m)}_{i,\ j} = 1$ for $i = 1,\ \ldots,\ n$. That is, the bWH(1) graduation is a linear smoother whose filter weights sum to unity.

(iii) Since $\boldsymbol{S}^{(m)}$ is centrosymmetric, it can be represented as in Abu-Jeib [27, Lemma 2.3]. Since $\boldsymbol{S}^{(m)}$ is also symmetric, the computational effort is further reduced. Thus, it is sufficient to compute about a quarter of the total entries.

(iv) Let $\boldsymbol{J}_n$ be the $n \times n$ exchange matrix. Explicitly, it is defined by $\boldsymbol{J}_n = [\boldsymbol{e}_n,\ \ldots,\ \boldsymbol{e}_1]$, where $\boldsymbol{e}_j$ denotes the j-th column of $\boldsymbol{I}_n$ for $j = 1,\ \ldots,\ n$. Then, we can provide another proof of Corollary 2 using $\boldsymbol{J}_n$, which is given in Appendix A. 2.

(v) A GNU Octave user-defined function to calculate $\widehat{\boldsymbol{x}}^{(m)}$ based on Proposition 2 and Corollary 2, `calc_xhat_bisymmetry`, is provided in Appendix B. 2.

4. Supplementary

In Section 1, we stated that (i) the bWH(1) graduation is suitable for smoothing data without a trend and (ii) the bHP filter developed by Phillips and Shi [4] is designed to recover the trend elements retained in the trend residuals. This section provides additional information about those issues, which give some reasons why the bWH(1) graduation is required. In Section 4.1, we present one of the reasons of the first issue. In Section 4.2, we illustrate how the bWH(1) graduation recovers the smooth elements retained in the residuals of the WH(1) graduation, $\boldsymbol{y} - \widetilde{\boldsymbol{x}}$.

4.1 *On the first issue*

Since $\phi_i = 1 - \left(\frac{\lambda q_i}{1+\lambda q_i}\right)^m \to 0$ as $\lambda \to \infty$ for $i = 2,\ \ldots,\ n$, from (10), it follows that

$$\boldsymbol{S}^{(m)} = \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top + \sum_{k=2}^{n} \phi_k \boldsymbol{p}_k \boldsymbol{p}_k^\top \to \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top, \quad (\lambda \to \infty). \tag{13}$$

In addition, given $\phi_1 = 1$, since $\phi_i = 1 - \left(\frac{\lambda q_i}{1+\lambda q_i}\right)^m \to 1$ as $\lambda \to 0$ for $i = 2, \ldots, n$, from (7), it follows that

$$\boldsymbol{S}^{(m)} = \boldsymbol{P}\boldsymbol{\Phi}\boldsymbol{P}^\top \to \boldsymbol{P}\boldsymbol{P}^\top = \boldsymbol{I}_n, \quad (\lambda \to 0). \tag{14}$$

Accordingly, we have the following results.

Proposition 3 (i) $\widehat{\boldsymbol{x}}^{(m)} = \boldsymbol{S}^{(m)}\boldsymbol{y} \to \boldsymbol{\iota}(\boldsymbol{\iota}^\top\boldsymbol{\iota})^{-1}\boldsymbol{\iota}^\top\boldsymbol{y} = \bar{y}\boldsymbol{\iota}$ as $\lambda \to \infty$ and (ii) $\widehat{\boldsymbol{x}}^{(m)} = \boldsymbol{S}^{(m)}\boldsymbol{y} \to \boldsymbol{y}$ as $\lambda \to 0$.

Proof. (i) and (ii) immediately follow from (13) and (14), respectively.

Remark 4 We give two remarks on Proposition 3.

(i) An $n \times n$ matrix $\boldsymbol{\iota}(\boldsymbol{\iota}^\top\boldsymbol{\iota})^{-1}\boldsymbol{\iota}^\top$ is the orthogonal projection matrix onto the space spanned by $\boldsymbol{\iota}$ and $\bar{y} = (\boldsymbol{\iota}^\top\boldsymbol{\iota})^{-1}\boldsymbol{\iota}^\top\boldsymbol{y} = \arg\min_{\mu\in} \|\boldsymbol{y} - \mu\boldsymbol{\iota}\|^2$.

(ii) The bHP filter does not have the property (i) in Proposition 3. For this reason, the bWH(1) graduation is more suitable for smoothing data without a trend than the bHP filter.

4.2 *On the second issue*

Since $\boldsymbol{S}^{(2)} = \boldsymbol{I}_n - (\boldsymbol{I}_n - \boldsymbol{S})^2 = \boldsymbol{I}_n - (\boldsymbol{I}_n - 2\boldsymbol{S} + \boldsymbol{S}^2) = \boldsymbol{S} + \boldsymbol{S}(\boldsymbol{I}_n - \boldsymbol{S})$, it follows that

$$\widehat{\boldsymbol{x}}^{(2)} = \boldsymbol{S}^{(2)}\boldsymbol{y} = \boldsymbol{S}\boldsymbol{y} + \boldsymbol{S}(\boldsymbol{I}_n - \boldsymbol{S})\boldsymbol{y} = \widetilde{\boldsymbol{x}} + \boldsymbol{S}(\boldsymbol{y} - \widetilde{\boldsymbol{x}}). \tag{15}$$

Given that $\boldsymbol{S} = (\boldsymbol{I}_n + \lambda\boldsymbol{D}^\top\boldsymbol{D})^{-1}$ is a low-pass filter, $\boldsymbol{S}(\boldsymbol{y} - \widetilde{\boldsymbol{x}})$ in (15) represents recovered smooth elements retained in the residuals of the WH(1) graduation. It is the gain of the boosting. Incidentally, such a gain occurs because $\boldsymbol{S}$ is not orthogonal to $\boldsymbol{y} - \widetilde{\boldsymbol{x}}$. This is in contrast to OLS regression, where the hat matrix is orthogonal to the residual vector.

5. Concluding remarks

In this paper, after defining the boosted WH graduation of order 1, we obtained some results on it. Specifically, we showed that it is a graph spectral filter using the DCT, and then provided a simple formula for the (i, j) entry of its smoother matrix. We also showed that it is a linear smoother such that the filter weights sum to unity and the smoother matrix is bisymmetric. They were summarized in Propositions 1-2 and Corollaries 1-2. In addition, we gave some reasons why the bWH(1) graduation is required and provided GNU Octave user-defined functions based on the results obtained.

Finally, we give a remark. To apply the bWH(1) graduation, the values of the two parameters in (3), m and λ, must be specified. To specify m in the bHP filter, Phillips and Shi [4] suggested using an information criterion, and we can consider a similar one:

$$IC(m) = \frac{\|\boldsymbol{y} - \widehat{\boldsymbol{x}}^{(m)}\|^2}{\|\boldsymbol{y} - \widetilde{\boldsymbol{x}}\|^2} + \log(n)\frac{\text{tr}(\boldsymbol{S}^{(m)})}{\text{tr}(\boldsymbol{I}_n - \boldsymbol{S})}. \tag{16}$$

Here, $\text{tr}(\boldsymbol{S}^{(m)})$, which equals $\text{tr}(\boldsymbol{\Phi}) = \sum_{i=1}^n \phi_i$, denotes the effective degrees of freedom Hastie et al. [28] of the bWH(1) graduation. Specifying these parameters is an important research topic, but is beyond the scope of this paper. We are researching this issue and will report our findings in the future.

Conflict of interest

The authors declare there is no conflict of interest at any point with reference to research findings.

References

[1] Bohlmann G. Ein Ausgleichungsproblem. *Nachrichten Von Der Gesellschaft Der Wissenschaften Zu Gottingen, Mathematisch-Physikalische Klasse*. 1899. p.260-271.

[2] Weinert HL. Efficient computation for Whittaker-Henderson smoothing. *Computational Statistics and Data Analysis*. 2007; 52(2): 959-974.

[3] Hodrick RJ, Prescott EC. Postwar U. S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*. 1997; 29(1): 1-16.

[4] Phillips PCB, Shi Z. Boosting: Why you can use the HP filter. *International Economic Review*. 2021; 62(2): 521-570.

[5] Knight K. *The Boosted Hodrick-Prescott Filter, Penalized Least Squares, and Bernstein Polynomials*. University of Toronto; 2021. Available from: https://utstat.utoronto.ca/keith/papers/hp-pls.pdf [Accessed 9th September 2023].

[6] Biswas E, Sabzikar F, Phillips PCB. *Boosting the Hp Filter for Trending Time Series With Long Range Dependence. Yale University EliScholar-A Digital Platform for Scholarly Publishing at Yale*; 2022. Available from: https://cowles.yale.edu/sites/default/files/2022-12/d2347.pdf [Accessed 15th July 2023].

[7] Hall V, Thomson P. *A Boosted Hp Filter for Business Cycle Analysis: Evidence From New Zealand's Small Open Economy*. CAMA Working Papers 2022-45; 2022.

[8] Mei ZW, Phillips PCB, Shi ZT. The boosted HP filter is more general than you might think. *arXiv:2209.09810*. 2022. Available from: https://arxiv.org/abs/2209.09810.

[9] Tomal M. Testing for overall and cluster convergence of housing rents using robust methodology: Evidence from Polish provincial capitals. *Empirical Economics*. 2022; 62(4): 2023-2055.

[10] Widiantoro DM. *Countercyclical Capital Buffer: Building the Resilience or Taming the Rapid Financial Cycle*? MPRA Paper 113507; 2022.

[11] Lu Y, Pagan A. *To Boost or Not to Boost? That Is the Question*. CAMA Working Paper 12/2023; 2023.

[12] Shuman DI, Narang SK, Frossard P, Ortega A, Vandergheynst P. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Processing Magazine*. 2013; 30(3): 83-98.

[13] Ahmed N, Natarajan T, Rao KR. Discrete cosine transform. *IEEE Transactions on Computers*. 1974; C-23(1): 90-93.

[14] Yamada H, Jahra FT. Explicit formulas for the smoother weights of the Whittaker-Henderson graduation of order 1. *Communications in Statistics-Theory and Methods*. 2018; 48(12): 3153-3161.

[15] Yamada H. A smoothing method that looks like the Hodrick-Prescott filter. *Econometric Theory*. 2020; 36(5): 961-981.

[16] Yamada, H. A note on Whittaker-Henderson graduation: Bisymmetry of the smoother matrix. *Communications in Statistics-Theory and Methods*. 2020; 49(7): 1629-1634.

[17] Nakatsukasa Y, Saito N, Woei E. Mysteries around the graph Laplacian eigenvalue 4. *Linear Algebra and its Applications*. 2013; 438(8): 3231-3246.

[18] Strang G, MacNamara S. Functions of difference matrices are Toeplitz plus Hankel. *SIAM Review*. 2014; 56(3): 525-546.

[19] Von Neumann J. Distribution of the ratio of the mean square successive difference to the variance. *Annals of Mathematical Statistics*. 1941; 12(4): 367-395.

[20] Strang G. The discrete cosine transform. *SIAM Review*. 1999; 41(1): 135-147.

[21] Yamada H. A pioneering study on discrete cosine transform. *Communications in Statistics-Theory and Methods*. 2022; 51(15): 5364-5368.

[22] Bapat RB. *Graphs and Matrices*. 2nd ed. London: Springer; 2014.

[23] Gallier J. Spectral theory of unsigned and signed graphs. Applications to graph clustering: A survey. *arXiv: 1601.04692*. 2016. Available from: https://arxiv.org/abs/1601.04692.

[24] Garcia D. Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational Statistics and Data Analysis*. 2010; 54(4): 1167-1178.

[25] Yamada H. Geary's *c* and spectral graph theory. *Mathematics*. 2021; 9(19): 2465.
[26] Yamada H. Geary's *c* and spectral graph theory: A complement. *Mathematics*. 2023; 11(20): 4228.
[27] Abu-Jeib IT. Centrosymmetric matrices: Properties and an alternative approach, *Canadian Mathematical Quarterly*. 2002; 10(4): 429-445.
[28] Hastie, T. Tibshirani R, Friedman J. *The Elements of Statistical Learning*. 2nd ed. New York: Springer; 2009.

Appendix A: Proofs

A.1 Proof of (6)

(i) Given $q_1 = 0$, $\phi_1 = 1$ immediately follows. (ii) Given that both λ and q_n are positive, it follows that $0 < \frac{\lambda q_n}{1+\lambda q_n} < 1$. Then, given that m is a positive integer, we obtain $0 < \left(\frac{\lambda q_n}{1+\lambda q_n}\right)^m < 1$, which yields $\phi_n = 1 - \left(\frac{\lambda q_n}{1+\lambda q_n}\right)^m > 0$. (iii) For $i = 2, \ldots, n$, it follows that $\phi_{i-1} - \phi_i = \left\{1 - \left(\frac{\lambda q_{i-1}}{1+\lambda q_{i-1}}\right)^m\right\} - \left\{1 - \left(\frac{\lambda q_i}{1+\lambda q_i}\right)^m\right\} = \left(\frac{\lambda q_i}{1+\lambda q_i}\right)^m - \left(\frac{\lambda q_{i-1}}{1+\lambda q_{i-1}}\right)^m$. Here, it follows that $\frac{\lambda q_i}{1+\lambda q_i} - \frac{\lambda q_{i-1}}{1+\lambda q_{i-1}} = \frac{\lambda q_i(1+\lambda q_{i-1}) - \lambda q_{i-1}(1+\lambda q_i)}{(1+\lambda q_i)(1+\lambda q_{i-1})} = \frac{\lambda(q_i - q_{i-1})}{(1+\lambda q_i)(1+\lambda q_{i-1})}$, from which, given that $\lambda > 0$ and $q_i > q_{i-1} \geq 0$, we obtain $\frac{\lambda q_i}{1+\lambda q_i} > \frac{\lambda q_{i-1}}{1+\lambda q_{i-1}} \geq 0$. Thus, given that m is a positive integer, it follows that $\phi_{i-1} - \phi_i = \left(\frac{\lambda q_i}{1+\lambda q_i}\right)^m - \left(\frac{\lambda q_{i-1}}{1+\lambda q_{i-1}}\right)^m > 0$, which yields $\phi_1 > \cdots > \phi_n$. Combining (i)-(iii) yields (6).

A.2 Another proof of Corollary 2

We prove that $\boldsymbol{S}^{(m)}$ is centrosymmetric by showing $\boldsymbol{J}_n\boldsymbol{S}^{(m)}\boldsymbol{J}_n = \boldsymbol{S}^{(m)}$. From Yamada [16, Theorem 2.2], it follows that $\boldsymbol{J}_n\boldsymbol{S}\boldsymbol{J}_n = \boldsymbol{S}$. In addition, $\boldsymbol{J}_n^2 = \boldsymbol{I}_n$. Then, $\boldsymbol{J}_n(\boldsymbol{I}_n - \boldsymbol{S})\boldsymbol{J}_n = \boldsymbol{J}_n^2 - \boldsymbol{J}_n\boldsymbol{S}\boldsymbol{J}_n = (\boldsymbol{I}_n - \boldsymbol{S})$. Suppose that $\boldsymbol{J}_n(\boldsymbol{I}_n - \boldsymbol{S})^{r-1}\boldsymbol{J}_n = (\boldsymbol{I}_n - \boldsymbol{S})^{r-1}$, where r is an integer greater than or equal to 2. Then, it follows that

$$\boldsymbol{J}_n(\boldsymbol{I}_n - \boldsymbol{S})^r\boldsymbol{J}_n = \boldsymbol{J}_n(\boldsymbol{I}_n - \boldsymbol{S})^{r-1}(\boldsymbol{I}_n - \boldsymbol{S})\boldsymbol{J}_n = \boldsymbol{J}_n(\boldsymbol{I}_n - \boldsymbol{S})^{r-1}\boldsymbol{J}_n\boldsymbol{J}_n(\boldsymbol{I}_n - \boldsymbol{S})\boldsymbol{J}_n$$

$$= (\boldsymbol{I}_n - \boldsymbol{S})^{r-1}(\boldsymbol{I}_n - \boldsymbol{S}) = (\boldsymbol{I}_n - \boldsymbol{S})^r.$$

Thus, by mathematical induction, $\boldsymbol{J}_n(\boldsymbol{I}_n - \boldsymbol{S})^m\boldsymbol{J}_n = (\boldsymbol{I}_n - \boldsymbol{S})^m$. Consequently,

$$\boldsymbol{J}_n\boldsymbol{S}^{(m)}\boldsymbol{J}_n = \boldsymbol{J}_n\{\boldsymbol{I} - (\boldsymbol{I}_n - \boldsymbol{S})^m\}\boldsymbol{J}_n = \boldsymbol{J}_n^2 - \boldsymbol{J}_n(\boldsymbol{I}_n - \boldsymbol{S})^m\boldsymbol{J}_n = \boldsymbol{I}_n - (\boldsymbol{I}_n - \boldsymbol{S})^m$$

$$= \boldsymbol{S}^{(m)}.$$

Thus, $\boldsymbol{S}^{(m)}$ is centrosymmetric. Therefore, given that $\boldsymbol{S}^{(m)}$ is symmetric, $\boldsymbol{S}^{(m)}$ is bisymmetric.

Appendix B: GNU octave user-defined functions

In this section, we provide some GNU Octave user-defined functions. Among them, `calc_xhat_dct` is a function to calculate $\widehat{\boldsymbol{x}}^{(m)}$ based on Proposition 1. It utilizes `dct` and `idct` functions. `calc_xhat_bisymmetry` is a function to calculate $\widehat{\boldsymbol{x}}^{(m)}$ based on Proposition 2 and Corollary 2.

B.1 calc_xhat_dct

```
function [xhat] = calc_xhat_dct (y, lambda, m)
    pkg load signal;
    n = length (y);
    [phivec] = calc_phivec (n, m, lambda);
    xhat = idct (phivec.* dct (y));
```

```
end
```

B.2 calc_xhat_bisymmetry

```
function [xhat] = calc_xhat_bisymmetry (y, lambda, m)
    n = length (y); Sm = zeros (n, n); Tm = zeros (n, n);
    I = eye (n); J = I (:, n: -1: 1);
    for i = 1: n
        for j = 1: n
            if i > = j
                if i + j < n + 1
                    Sm(i, j) = calc_sm_ij (n, m, lambda, i, j);
                elseif i + j == n + 1
                    Tm (i, j) = calc_sm_ij (n, m, lambda, i, j);
                end
            end
        end
    end
    Sm = Sm + Sm'-diag (diag (Sm));
    Sm = (Sm + J * Sm * J) + (Tm + J * Tm * J);
    xhat = Sm * y;
end
```

B.3 Subroutines

```
function [phivec] = calc_phivec (n, m, lambda )
    iota = ones (n, 1) ; tau = (1: n)';
    lamqvec = 2* lambda *(iota-cos ((tau-iota) * pi/n));
    phivec = iota-(lamqvec./(iota + lamqvec)).^m;
end
```

```
function [sm_ij ] = calc_sm_ij (n, m, lambda, i, j)
    [phivec] = calc_phivec (n, m, lambda);
    theta_i = (i-1/2) * pi/n; theta_j =(j-1/2)* pi/n;
    psi = 0;
    for k = 2: n
        psi = psi+phivec (k)* cos ((k-1)* theta_i )* cos ((k-1)* theta_j);
    end
    sm_ij = (1/n) +(2/n)* psi;
end
```