Research Article

Boualem Djehiche, Astrid Hilbert* and Hiba Nassar

On the functional Hodrick–Prescott filter with non-compact operators

DOI: 10.1515/rose-2016-0003
Received December 10, 2014; accepted August 15, 2015

Abstract: We study a version of the functional Hodrick–Prescott filter in the case when the associated operator is not necessarily compact but merely closed and densely defined with closed range. We show that the associated optimal smoothing operator preserves the structure obtained in the compact case when the underlying distribution of the data is Gaussian.

Keywords: Inverse problems, adaptive estimation, Hodrick–Prescott filter, smoothing, trend extraction, Gaussian measures on a Hilbert space

MSC 2010: 62G05, 62G20

Communicated by: Sergio Albeverio

1 Introduction

The study of functional data analysis is motivated by its applications in various fields of statistical estimations and statistical inverse problems (see Ramsay and Silverman [18], Bosq [1], Müller and Stadtmüller [17] and the references therein). One of the most common assumptions in the studies of statistical inverse problems is to deal with compact operators. This is due to their tractable spectral properties. The functional Hodrick–Prescott filter is often formulated as a statistical inverse problem that reconstructs an "optimal smooth signal" y that solves an equation $Ay = v$ corrupted by a noise v, which is a priori unobservable, from observations x corrupted by a noise u, which is also a priori unobservable, i.e.,

$$\begin{cases} x = y + u, \\ Ay = v, \end{cases} \tag{1.1}$$

where $A : H_1 \to H_2$ is a compact operator between two appropriate Hilbert spaces H_1 and H_2.

By introducing a smoothing operator B, the "optimal smooth signal" $y(B, x)$ associated with x is defined by

$$y(B, x) := \arg\min_y \{\|x - y\|^2_{H_1} + \langle Ay, BAy\rangle_{H_2}\} \tag{1.2}$$

provided that

$$\langle Ah, BAh\rangle_{H_2} \geq 0, \quad h \in H_1. \tag{1.3}$$

In [5], the optimal smoothing operator is characterized as the minimizer of the difference between the optimal smoothing signal and the best predictor $E[y|x]$ of the signal given the data x, when the noise u and the sig-

Boualem Djehiche: Department of Mathematics, Royal Institute of Technology, 100 44 Stockholm, Sweden, e-mail: boualem@math.kth.se
***Corresponding author: Astrid Hilbert:** School of Computer Science, Physics and Mathematics, Linnaeus University, Vejdesplats 7, 351 95 Växjö, Sweden, e-mail: astrid.hilbert@lnu.se
Hiba Nassar: School of Computer Science, Physics and Mathematics, Linnaeus University, Vejdesplats 7, 351 95 Växjö, Sweden, e-mail: hiba.nassar@lnu.se

nal v are independent Hilbert space-valued Gaussian random variables with zero means and covariance operators Σ_u and Σ_v. For more details on the classical Hodrick–Prescott filter, the reader is referred to [3, 4, 9, 21].

In this paper, we extend the functional Hodrick–Prescott filter to the case where the operator A is not necessarily compact. Moreover, we show that the optimal smoothing parameter preserves the structure obtained in [5] for the compact case.

An important class of non-compact operators to which we wish to extend the Hodrick–Prescott filter includes the Laplace operator

$$A = -\frac{d^2}{dt^2}$$

with Dirichlet boundary conditions and with domain

$$D(A) = \{y \in H^2([0,1]) : y(0) = y(1) = 0\},$$

where $H^2([0,1])$ is the Sobolev space of functions whose weak derivatives of order less than or equal to two belong to $L^2([0,1])$ (see, e.g., [2] for further examples).

The paper is organized as follows. In Section 2, we generalize the functional Hodrick–Prescott filter under the assumption that the operator A is closed and densely defined with closed range. In Section 3, we prove that the optimal smoothing operator maintains the same form when the covariance operators Σ_u and Σ_v are trace class operators. In Section 4, we illustrate this filter with two examples. Finally, in Section 5, we extend this characterization to the case where the covariance operators Σ_u and Σ_v are not trace class such as, e.g., white noise.

2 A functional Hodrick–Prescott filter with closed operator

Let H_1 and H_2 be two separable Hilbert spaces with norms $\|\cdot\|_{H_i}$ and inner products $\langle\cdot,\cdot\rangle_{H_i}$, $i = 1, 2$, and let $x \in H_1$ be a functional time series of observables. Given a linear operator $A : H_1 \to H_2$, the Hodrick–Prescott filter extracts an "optimal smooth signal" $y \in H_1$ that solves an equation $Ay = v$ corrupted by a noise v, which is a priori unobservable, from observations x corrupted by a noise u, which is also a priori unobservable, i.e.,

$$\begin{cases} x = y + u, \\ Ay = v. \end{cases}$$

The optimality of the extracted signal is achieved by the following Tikhonov–Phillips regularization of (1.1) by introducing a linear operator $B : H_2 \to H_2$ which acts as a smoothing parameter, i.e.,

$$y(B,x) := \arg\min_y \{\|x - y\|_{H_1}^2 + \langle Ay, BAy\rangle_{H_2}\}$$

provided that

$$\langle Ah, BAh\rangle_{H_2} \geq 0, \quad h \in H_1.$$

As suggested in [5], the optimal smoothing operator minimizes the gap between the conditional expected value $E[y|x]$ of y given x, which is the best predictor of y given x, and $y(B,x)$, i.e.,

$$\hat{B} = \arg\min_B \|E[y|x] - y(B,x)\|_{H_1}^2. \tag{2.1}$$

The main purpose of this work is to extend the characterization of the optimal smoothing operator obtained in [5] to the case where the linear operator A is not necessarily compact and u and v are independent Hilbert space-valued generalized Gaussian random variables with zero means and covariance operators Σ_u and Σ_v.

Assumption 1. *We assume for the linear operator $A : H_1 \to H_2$ that*
(i) *it is closed and defined on a dense subspace $\mathcal{D}(A)$ of H_1;*
(ii) *its range* $\operatorname{Ran}(A)$ *is closed.*

Its Moore–Penrose generalized inverse $A^\dagger$ is defined on

$$\mathcal{D}(A^\dagger) = \operatorname{Ran}(A) + \operatorname{Ran}(A)^\perp,$$

i.e., $\operatorname{Ker}(A^\dagger) = \operatorname{Ran}(A)^\perp$. Note that Assumption 1 is equivalent to the fact that $A^\dagger$ is bounded (see [7, 8, 11, 14]).

For all $v \in \mathcal{D}(A^\dagger)$, the set of all solutions of the equation

$$Ay = v, \quad y \in \mathcal{D}(A),$$

is given by

$$\{y^\dagger + y_0 : y_0 \in \operatorname{Ker}(A)\},$$

where $y^\dagger$ is the unique minimal-norm solution given by $y^\dagger = A^\dagger v$ (see [13]). Hence, for arbitrary $y_0 \in \operatorname{Ker}(A)$, we have

$$y = y_0 + A^\dagger v \tag{2.2}$$

and, in view of (1.1),

$$x = y_0 + A^\dagger v + u. \tag{2.3}$$

Let $\Pi := A^\dagger A$. By the Moore–Penrose equations, we have $\Pi A^\dagger = A^\dagger$, $\Pi^2 = \Pi$ and $\Pi^* = \Pi$ (self-adjoint). Therefore, Π is an orthogonal projector. It is easily checked that, for every $\xi \in H_1$, the elements $\Pi\xi$ and $(I_{H_1} - \Pi)\xi$ are orthogonal, i.e.,

$$\langle \Pi\xi, (I_{H_1} - \Pi)\xi\rangle = 0$$

and

$$\operatorname{Ker}(A) = \operatorname{Ker}(\Pi) = \operatorname{Ran}(I_{H_1} - \Pi).$$

Moreover, we have $(I_{H_1} - \Pi)y = y_0$ and $A^\dagger v = \Pi y$.

In the next proposition, we show that (1.2) has a unique solution for a class of linear smoothing operators B satisfying (1.3).

Proposition 2.1. *Let $A : H_1 \to H_2$ be a closed, linear operator such that its domain is dense in H_1. Assume further that the smoothing operator $B : H_2 \to H_2$ is closed, densely defined and satisfies*

$$\langle Ah, BAh\rangle_{H_2} \geq 0, \quad h \in H_1. \tag{2.4}$$

Then, there exists a unique $y(B, x) \in H_1$ which minimizes the functional

$$J_B(y) = \|x - y\|^2_{H_1} + \langle Ay, BAy\rangle_{H_2}.$$

This minimizer is given by the formula

$$y(B, x) = (I_{H_1} + A^* BA)^{-1} x. \tag{2.5}$$

Proof. It is immediate to check that the minimizer of the functional $\|x - y\|^2_{H_1} + \langle Ay, BAy\rangle_{H_2}$ is $(I_H + A^*BA)^{-1}x$ (see [10]) provided that the function $(I_H + A^*BA)^{-1}$ exists everywhere. But, since the operator $D := \sqrt{B}A$ is closed and densely defined, thanks to a result by Neumann (see [19, Chapter VII, Section 118]), we have that $(I_H + A^*BA)^{-1} = (I + D^*D)^{-1}$ exists everywhere and is bounded. □

3 Hodrick–Prescott filter associated with trace class covariance operators

In this section, we prove that the optimal smoothing operator which solves (2.1) has the same structure as in [5], when u and v are independent Gaussian random variables with zero mean and trace class covariance operators Σ_u and Σ_v (see [22]).

In view of (2.2) and (2.3), a stochastic model for (x, y) being determined by models for y_0 and (u, v), we assume the following.

Assumption 2. *y_0 deterministic.*

Assumption 3. *u and v are independent Gaussian random variables with zero mean and covariance operators Σ_u and Σ_v respectively.*

Assumption 2 is made to ease the analysis. The independence between u and v imposed in Assumption 3 is natural because a priori there should not be any dependence between the "residual" u which is due to the noisy observation x and the required degree of smoothness of the signal y.

Assumption 3 implies that $\Pi y = A^\dagger v$ and u are also independent. Thus, with regard to the decomposition

$$x = y_0 + \Pi y + \Pi u + (I_{H_1} - \Pi)u,$$

it is natural to assume that even the orthogonal random variables Πu and $(I_{H_1} - \Pi)u$ are independent. This would mean that the input x is decomposed into three independent random variables. This is actually the case for the classical Hodrick–Prescott filter. Also, as we will show below, thanks to this property, the optimal smoothing operator has the form of a "noise-to-signal ratio" in line with the classical Hodrick–Prescott filter.

Assumption 4. *The orthogonal (in H_1) random variables Πu and $(I_{H_1} - \Pi)u$ are independent, i.e.,*

$$\Pi\Sigma_u = \Sigma_u\Pi. \tag{3.1}$$

We note that (3.1) is equivalent to

$$\Pi\Sigma_u\Pi = \Pi\Sigma_u.$$

Given Assumptions 2 and 3, by (2.2) and (2.3), it holds that (x, y) is Gaussian with mean $(E[x], E[y]) = (y_0, y_0)$ and covariance operator

$$\Sigma = \begin{pmatrix} \Sigma_u + Q_v & Q_v \\ Q_v & Q_v \end{pmatrix},$$

where

$$Q_v := A^\dagger\Sigma_v(A^\dagger)^*.$$

Lemma 3.1. *The linear operator Q_v is trace class. Moreover, the linear operator*

$$T := Q_v[\Sigma_u + Q_v]^{-1/2}$$

is Hilbert–Schmidt.

Proof. Since the covariance operator Σ_v is a trace class operator, $\Sigma_v^{1/2}$ is Hilbert–Schmidt. Therefore, $A^\dagger\Sigma_v^{1/2}$ is Hilbert–Schmidt, since, by Assumption 1, $A^\dagger$ is bounded. Hence,

$$Q_v = A^\dagger\Sigma_v^{1/2}(A^\dagger\Sigma_v^{1/2})^*$$

is a trace class operator since it is a product of two Hilbert–Schmidt operators. Furthermore, since $\Sigma_u + Q_v$ is injective and trace-class, the operator $[\Sigma_u + Q_v]^{1/2}$ is Hilbert–Schmidt. Hence, $T := Q_v[\Sigma_u + Q_v]^{-1/2}$ is Hilbert–Schmidt. □

We may apply [16, Theorem 2] to obtain the conditional expectation of the signal y given the functional data x, i.e.,

$$E[y|x] = y_0 + Q_v[\Sigma_u + Q_v]^{-1}(x - y_0).$$

The following theorem is a generalization of [5, Theorem 4].

Theorem 3.2. *Under Assumptions 1–4, the smoothing operator*

$$\hat{B} := (A^\dagger)^*\Sigma_u A^*\Sigma_v^{-1} \tag{3.2}$$

is the unique operator which satisfies

$$\hat{B} = \arg\min_{B} \|E[y|x] - y(B, x)\|_{H_1},$$

where the minimum is taken with respect to all linear, closed and densely defined operators B which satisfy the positivity condition (2.4).

Proof. The proof is similar to that of [5, Theorem 4]. □

4 Examples

In this section, we apply Theorem 3 to two examples for which the operators are densely defined with closed range.

Example 4.1. Inspired by an example discussed in [12], we consider the operator $A : l^2 \to l^2$ defined by

$$A(x_1, x_2, x_3, \dots, x_n, \dots) = (0, 2x_2, 3x_3, \dots, nx_n, \dots)$$

with domain

$$\mathcal{D}(A) = \left\{ x := (x_1, x_2, x_3, \dots, x_n, \dots) \in l^2 : \sum_{j=1}^{\infty} |jx_j|^2 < \infty \right\}.$$

This operator is self-adjoint, unbounded, closed and densely defined with $\overline{\mathcal{D}(A)} = l^2$.

Now consider the Hodrick–Prescott filter associated with the operator A under the assumption that u and v are independent Gaussian random variables with zero means and covariance operators of trace class of the form

$$\Sigma_u x = (\sigma_1^u x_1, \sigma_2^u x_2, \sigma_3^u x_3, \dots, \sigma_n^u x_n, \dots)$$

and

$$\Sigma_v x = (\sigma_1^v x_1, \sigma_2^v x_2, \sigma_3^v x_3, \dots, \sigma_n^v x_n, \dots),$$

respectively.

In view of the form of the operator A, an appropriate class of smoothing operators B is

$$B(x_1, x_2, x_3, \dots, x_n, \dots) = (b_1 x_1, b_2 x_2, b_3 x_3, \dots, b_n x_n, \dots),$$

where the coefficients b_n, $n = 1, 2, \dots$, are chosen so that the operator B is closed, densely defined and satisfies the positivity condition (2.4). In view of Theorem 3.2, the optimal smoothing operator B given by (3.2) reads

$$\hat{B}x = A^{-1}\Sigma_u A \Sigma_v^{-1} x = \left(0, \frac{\sigma_2^u}{\sigma_2^v} x_2, \frac{\sigma_3^u}{\sigma_3^v} x_3, \dots, \frac{\sigma_n^u}{\sigma_n^v} x_n, \dots\right), \quad x \in l^2.$$

Moreover, the corresponding optimal signal given by (2.5) is

$$y(\hat{B}, x) = (I_{H_1} + A^* \hat{B} A)^{-1} x = \left(x_1, \frac{1}{4\hat{b}_2 + 1} x_2, \frac{1}{9\hat{b}_3 + 1} x_3, \dots, \frac{1}{n^2 \hat{b}_n + 1} x_n, \dots\right),$$

where

$$\hat{b}_j := \frac{\sigma_j^u}{\sigma_j^v}, \quad j = 1, 2, \dots .$$

Example 4.2. Consider the Laplace operator

$$A = -\frac{d^2}{dt^2}$$

with Dirichlet boundary conditions and with domain

$$D(A) = \{y \in H^2([0, 1]) : y(0) = y(1) = 0\},$$

where $H^2([0, 1])$ is the Sobolev space of functions whose weak derivatives of order less than or equal to two belong to $L^2([0, 1])$.

The signal process y corrupted by v satisfies

$$Ay(t) = -\frac{d^2y(t)}{dt^2} = v(t), \quad y \in D(A), \tag{4.1}$$

The Laplacian A is a one-to-one, non-negative, self-adjoint, closed and unbounded operator with domain $D(A)$ dense in $L^2([0,1])$ (see, e.g., [2]). The eigenvalues and eigenvectors of A satisfy

$$\begin{cases} \lambda_n = n^2\pi^2, & n \geq 1,\ n \in \mathbb{N}, \\ e_n(t) = \sqrt{2}\sin n\pi t. \end{cases}$$

The inverse of the operator A is a self-adjoint Hilbert–Schmidt operator given by

$$(A^{-1}x)(t) = \int_0^1 G(t,s)x(s)\,ds,$$

where the Green's function $G : [0,1] \times [0,1] \to [0,1]$ is given by

$$G(t,s) := \begin{cases} (1-t)s, & 0 \leq s \leq t, \\ t(1-s), & t \leq s \leq 1. \end{cases}$$

Hence, the operator A can be written as

$$Ay(t) = \sum_{n=1}^{\infty} n^2\pi^2 \langle y, e_n\rangle e_n(t)$$

and the solution of (4.1) is given in terms of eigenvalues and eigenvectors by

$$y(t) = A^{-1}v(t) = \sum_{n=1}^{\infty} \frac{1}{n^2\pi^2}\langle v, e_n\rangle e_n(t).$$

Now, consider the Hodrick–Prescott filter associated with the operator A under the assumption that u and v are independent Gaussian random variables with zero means and covariance operators of trace class of the form

$$\Sigma_u h(t) = \sum_{n=1}^{\infty} \sigma_n^u \langle h, e_n\rangle e_n(t)$$

and

$$\Sigma_v h(t) = \sum_{n=1}^{\infty} \sigma_n^v \langle h, e_n\rangle e_n(t),$$

respectively, where the sums converge in the operator norm.

The smoothing operator B is defined as

$$Bh(t) = \sum_{n=1}^{\infty} \beta_n \langle h, e_n\rangle e_n(t),$$

where the coefficients β_n, $n = 1, 2, \ldots$, are chosen so that the operator B is closed, densely defined and satisfies the positivity condition (2.4). By Theorem 3.2, the optimal smoothing operator B given by (3.2) reads

$$\hat{B}h(t) = A^{-1}\Sigma_u A \Sigma_v^{-1} h(t) = \sum_{n=1}^{\infty} \frac{\sigma_n^u}{\sigma_n^v}\langle h, e_n\rangle e_n(t).$$

The corresponding optimal signal given by (2.5) is

$$y(\hat{B}, x) = (I_{L^2(0,1)} + A^*\hat{B}A)^{-1}x = \sum_{n=1}^{\infty}\left(1 + n^4\pi^4\frac{\sigma_n^u}{\sigma_n^v}\right)^{-1}\langle x, e_n\rangle e_n(t).$$

5 Extension to non-trace class covariance operators

In this section, we show that the characterization (3.2) of the optimal smoothing operator is preserved even when the covariance operators of u and v are not necessarily trace class operators.

We assume that $u \sim N(0, \Sigma_u)$ and $v \sim N(0, \Sigma_v)$, where Σ_u and Σ_v are self-adjoint, positive-definite and bounded, but not trace class operators on H_1 and H_2, respectively. One important case of this extension is when u and v are white noise with covariance operators of the form $\Sigma_u = \sigma_u I_{H_1}$ and $\Sigma_v = \sigma_v I_{H_2}$, respectively, for some constants σ_u and σ_v.

Following Rozanov [20] (see also Lehtinen, Päivärinta and Somersalo [15]), we consider these Gaussian variables as generalized random variables on an appropriate Hilbert scale (or nuclear countable Hilbert space), where the covariance operators can be maximally extended to self-adjoint, positive-definite, bounded and trace class operators on an appropriate domain.

We first construct the Hilbert scale appropriate to our setting. This is performed using the linear operator A as follows (see Engl, Hanke and Neubauer [6] for further details).

In view of Assumption 1, the operator $A^\dagger : H_2 \to H_1$ is linear and bounded. Putting $H_3 := \operatorname{Ran} A$, then H_3 is a Hilbert space since it is a closed subspace of the Hilbert space H_2. Let $\bar{A}^\dagger$ be the restriction of $A^\dagger$ on H_3, i.e., $\bar{A}^\dagger : H_3 \to H_1$. Then, $\bar{A}^\dagger$ is an injective bounded linear operator.

Remark 5.1. In view of the Hodrick–Prescott filter (1.1), we have $v \in \operatorname{Ran}(A) = H_3$, i.e., it can be seen as an H_3-random variable with covariance operator $\Sigma_v : H_3 \to H_3$.

Set

$$K_1 := (\bar{A}^\dagger(\bar{A}^\dagger)^*)^{-1} : H_1 \to H_1.$$

We can define the fractional power of the operator K_1 by

$$K_1^s h = (\bar{A}^\dagger(\bar{A}^\dagger)^*)^{-s} h, \quad h \in H_1,\ s \geq 0,$$

and we define its domain by

$$\mathcal{D}(K_1^s) := \{h \in H_1 : (\bar{A}^\dagger(\bar{A}^\dagger)^*)^{-s} h \in H_1\}.$$

Let $\mathcal{M}$ be the set of all elements x for which all the positive integer powers of K_1 are defined, i.e.,

$$\mathcal{M} := \bigcap_{n=0}^{\infty} \mathcal{D}(K_1^n).$$

For $s \geq 0$, let H_1^s be the completion of $\mathcal{M}$ with respect to the Hilbert space norm induced by the inner product

$$\langle x, y\rangle_{H_1^s} := \langle K_1^s x, K_1^s y\rangle_{H_1}, \quad x, y \in \mathcal{M},$$

and let $H_1^{-s} := (H_1^s)^*$ denote the dual of H_1^s equipped with the inner product

$$\langle x, y\rangle_{H_1^{-s}} := \langle K_1^{-s} x, K_1^{-s} y\rangle_{H_1}, \quad x, y \in \mathcal{M}.$$

Then, $(H_1^s)_{s\in\mathbb{R}}$ is the Hilbert scale induced by the operator K_1.

The operator

$$K_2 := ((\bar{A}^\dagger)^* \bar{A}^\dagger)^{-1} : H_3 \to H_3$$

has the same properties as K_1. Repeating the same procedure as before, we get that $(H_3^s)_{s\in\mathbb{R}}$ is the Hilbert scale induced by the operator K_2, where the norm in H_3^n is given by $\|h\|_{H_3^n} = \|K_2^n h\|_{H_3}$, $h \in H_3^n$.

Noting that

$$H_1^{-n} = \operatorname{Im}((\bar{A}^\dagger(\bar{A}^\dagger)^*)^n) = (\bar{A}^\dagger(\bar{A}^\dagger)^*)^n(H_1),$$
$$H_3^{-n} = \operatorname{Im}(((\bar{A}^\dagger)^*\bar{A}^\dagger)^n) = ((\bar{A}^\dagger)^*\bar{A}^\dagger)^n(H_3)$$

with $\ker((\bar{A}^\dagger(\bar{A}^\dagger)^*)^n) = \ker(\bar{A}^\dagger) = \{0\}$ and $\ker(((\bar{A}^\dagger)^*\bar{A}^\dagger)^n) = \ker((\bar{A}^\dagger)^*) = \ker(\bar{A}) = \ker(A)$, it follows that the operator $\bar{A}^\dagger$ extends to a continuous operator from H_3^{-n} into H_1^{-n}, the operator $\bar{A} = A$ extends to a continuous operator from H_1^{-n} into H_3^{-n} and the operators $\bar{A}^\dagger(\bar{A}^\dagger)^*$ and $(\bar{A}^\dagger)^*\bar{A}^\dagger$ extend as well to continuous operators onto H_1^{-n} and H_3^{-n}, respectively.

Extending the Hodrick–Prescott filter to the larger Hilbert spaces H_1^{-n} and H_3^{-n} due to the flexibility offered by the Hilbert scale, where n is chosen so that the second moments $E[\|x\|^2_{H_1^{-n}}]$, $E[\|y\|^2_{H_1^{-n}}]$, $E[\|u\|^2_{H_1^{-n}}]$ and $E[\|v\|^2_{H_3^{-n}}]$ of the Gaussian random variables x, y, u and v in H_1^{-n} and H_3^{-n}, respectively, are finite. This amounts to making their respective covariance operators

$$\tilde{\Sigma}_u = (\bar{A}^\dagger(\bar{A}^\dagger)^*)^n \Sigma_u (\bar{A}^\dagger(\bar{A}^\dagger)^*)^n,$$
$$\tilde{\Sigma}_v = ((\bar{A}^\dagger)^*\bar{A}^\dagger)^n \Sigma_v ((\bar{A}^\dagger)^*\bar{A}^\dagger)^n$$

and

$$\tilde{\Sigma} = \begin{pmatrix} \tilde{\Sigma}_u + \tilde{Q}_v & \tilde{Q}_v \\ \tilde{Q}_v & \tilde{Q}_v \end{pmatrix},$$

where the operator

$$\tilde{Q}_v := \bar{A}^\dagger \tilde{\Sigma}_v (\bar{A}^\dagger)^*$$

is trace class.

We make the following assumption.

Assumption 5. *There is $n_0 > 0$ such that the covariance operators $\tilde{\Sigma}_u$, $\tilde{\Sigma}$ and $\tilde{\Sigma}_v$ are trace class on the Hilbert spaces H_1^{-n} and H_3^{-n}, respectively.*

It is worth noting that since $y_0 \in \ker(A) = \ker(((\bar{A}^\dagger)^*\bar{A}^\dagger)^n)$, then $\|y_0\|_{H_1^{-n}} = \|(\bar{A}^\dagger(\bar{A}^\dagger)^*)^n y_0\|_{H_1} = 0$. Hence, the $H_1^{-n} \times H_1^{-n}$-valued random vector (x, y) has mean $(E[x], E[y]) = (0, 0)$.

Summing up, by Assumption 5, for $n \geq n_0$, the vector (x, y) is an $H_1^{-n} \times H_1^{-n}$-valued Gaussian vector with mean $(0, 0)$ and covariance operator $\tilde{\Sigma}$. Thus, by [16, Theorem 2], we have

$$E[y|x] = \tilde{Q}_v[\tilde{\Sigma}_u + \tilde{Q}_v]^{-1} x \quad \text{a.s. in } H_1^{-n}$$

provided that the operator

$$\tilde{T} := \tilde{\Sigma}_{XY} \tilde{\Sigma}_X^{-1/2}$$

is Hilbert–Schmidt. But, in view of Assumption 5 and Lemma (3.1), the operator T is Hilbert–Schmidt.

The deterministic optimal signal associated with x in H_1^{-n}, $n \geq n_0$, is given by (cf. Proposition 2.1)

$$y(B, x) = (I_{H_1^{-n}} + A^*BA)^{-1} x,$$

which is the unique minimizer of the functional

$$J_B(y) = \|x - y\|^2_{H_1^{-n}} + \langle Ay, BAy \rangle_{H_3^{-n}}$$

with a linear operator $B : H_3^{-n} \to H_3^{-n}$ such that $\langle Ah, BAh \rangle_{H_3^{-n}} \geq 0$ for all $h \in H_1^{-n}$.

The following theorem gives an explicit expression of the optimal smoothing operator $\hat{B}$.

Theorem 5.2. *Let Assumption 5 hold. Then, the unique optimal smoothing operator associated with the Hodrick–Prescott filter associated with H_1^{-n}-valued data x is given by*

$$\hat{B}h := (\bar{A}^\dagger)^* \tilde{\Sigma}_u A^* \tilde{\Sigma}_v^{-1} h, \quad h \in H_3^{-n}. \tag{5.1}$$

Proof. The proof is similar to that of [5, Theorem 6]. □

5.1 The white noise case: Optimality of the noise-to-signal ratio

In this section, we show that the optimal smoothing operator $\hat{B}$ given by (5.1) reduces to the noise-to-signal ratio, where u and v are white noises. Assume that u and v are independent and Gaussian random variables with zero means and covariance operators $\Sigma_u = \sigma_u I_{H_1}$ and $\Sigma_v = \sigma_v I_{H_3}$, where I_{H_1} and I_{H_3} denote the H_1 and H_3 identity operators, respectively, and σ_u and σ_v are constant scalars. Assumption 5 reduces to the following assumption.

Assumption 6. *There is an $n_0 > 0$ such that $(\bar{A}^\dagger(\bar{A}^\dagger)^*)^{2n}$ and $((\bar{A}^\dagger)^*\bar{A}^\dagger)^{2n}$ are trace class for all $n \geq n_0$.*

Under this assumption, the associated covariance operators

$$\tilde{\Sigma}_u = (\bar{A}^\dagger(\bar{A}^\dagger)^*)^n \Sigma_u (\bar{A}^\dagger(\bar{A}^\dagger)^*)^n = \sigma_u(\bar{A}^\dagger(\bar{A}^\dagger)^*)^{2n},$$
$$\tilde{\Sigma}_v = ((\bar{A}^\dagger)^*\bar{A}^\dagger)^n \Sigma_v ((\bar{A}^\dagger)^*\bar{A}^\dagger)^n = \sigma_v((\bar{A}^\dagger)^*\bar{A}^\dagger)^{2n}$$

and

$$\tilde{Q}_v = \sigma_v A^\dagger((A^\dagger)^*A^\dagger)^{2n}(A^\dagger)^* = \sigma_v(A^\dagger(A^\dagger)^*)^{2n+1}$$

are trace class and (5.1), which gives the optimal smoothing operator $\hat{B}$, reduces to

$$\hat{B} = (\bar{A}^\dagger)^*\tilde{\Sigma}_u A^* \tilde{\Sigma}_v^{-1} h = \frac{\sigma_u}{\sigma_v} I_{H_3^{-n}},$$

i.e., $\hat{B}$ is the noise-to-signal ratio which is in the same pattern as in the classical Hodrick–Prescott filter.

Funding: The first author gratefully acknowledges the financial support of the Swedish Export Credit Corporation (SEK).

References

[1] D. Bosq, *Linear Processes in Function Spaces. Theory and Applications*, Lecture Notes in Statist. 149, Springer, New York, 2000.

[2] R. Dautray and J. L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology*, Springer, New York, 1988–1993.

[3] A. Dermoune, B. Djehiche and N. Rahmania, Consistent estimator of the smoothing parameter in the Hodrick–Prescott filter, *J. Japan Statist. Soc.* **38** (2008), no. 2, 225–241.

[4] A. Dermoune, B. Djehiche and N. Rahmania, Multivariate extension of the Hodrick–Prescott filter – Optimality and characterization, *Stud. Nonlinear Dyn. Econom.* **13** (2009), Article no. 4.

[5] B. Djehiche and H. Nassar, A functional Hodrick–Prescott filter, preprint (2013), http://arxiv.org/abs/1312.4936.

[6] H. W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problems*, Math. Appl. 375, Kluwer, Dordrecht, 1996.

[7] C. W. Groetsch, Spectral methods for linear inverse problems with unbounded operators, *J. Approx. Theory* **70** (1992), no. 1, 16–28.

[8] M. R. Hestenes, Relative self-adjoint operators in Hilbert space, *Pacific J. Math.* **11** (1961), 1315–1357.

[9] R. Hodrick and E. C. Prescott, Postwar U.S. business cycles: An empirical investigation, *J. Money Credit Bank.* **29** (1997), no. 1, 1–16.

[10] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, Appl. Math. Sci. 160, Springer, New York, 2004.

[11] T. Kato, *Perturbation Theory for Linear Operators*, 2nd ed., Grundlehren Math. Wiss. 132, Springer, Berlin, 1976.

[12] S. H. Kulkarni, M. T. Nair and G. Ramesh, Some properties of unbounded operators with closed range, *Proc. Indian Acad. Sci. Math. Sci.* **118** (2008), no. 4, 613–625.

[13] S. H. Kulkarni and G. Ramesh, Projection methods for computing Moore–Penrose inverses of unbounded operators, *Indian J. Pure Appl. Math.* **41** (2010), no. 5, 647–662.

[14] L. J. Lardy, A series representation for the generalized inverse of a closed linear operator, *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)* **58** (1975), no. 2, 152–157.

[15] M. S. Lehtinen, L. Päivärinta and E. Somersalo, Linear inverse problems for generalised random variables, *Inverse Problems* **5** (1989), no. 4, 599–612.

[16] A. Mandelbaum, Linear estimators and measurable linear transformations on a Hilbert space, *Z. Wahrsch. Verw. Gebiete* **65** (1984), no. 3, 385–397.

[17] H.-G. Müller and U. Stadtmüller, Generalized functional linear models, *Ann. Statist.* **33** (2005), no. 2, 774–805.

[18] J. O. Ramsay and B. W. Silverman, *Functional Data Analysis*, Springer, New York, 1997.
[19] F. Riesz and B. S. Nagy, *Functional Analysis*, Dover, New York, 1990.
[20] J. A. Rozanov, *Infinite-Dimensional Gaussian Distributions*, American Mathematical Society, Providence, 1971.
[21] E. Schlicht, Estimating the smoothing parameter in the so-called Hodrick–Prescott filter, *J. Japan Statist. Soc.* **35** (2005), no. 1, 99–119.
[22] A. V. Skorohod, *Integration in Hilbert Space*, Ergeb. Math. Grenzgeb. 79, Springer, Berlin, 1974.