Non-orthogonal Hilbert Projections in Trend Regression*

Peter C. B. Phillips
Cowles Foundation for Research in Economics
Yale University,
University of Auckland & University of York

and

Yixiao Sun
Department of Economics
Yale University

8 February 2001

Abstract

We consider detrending procedures in continuous time when the errors follow a Brownian motion and a diffusion process, respectively. We show that more efficient trend extraction is accomplished by non-orthogonal Hilbert projections in both cases.

JEL Classification: C22

Key words and phrases: Efficent Detrending, Brownian Motion, Diffusion Process, Hilbert Projection.

*Phillips thanks the NSF for research support under Grant No. SBR 97-30295. Sun thanks the Cowles Foundation for support under a Cowles Prize Fellowship. The paper was typed by the authors in SW2.5.

1. Problem

An observed continuous time process $X(t)$ is generated by the linear system

$$X(t) = \beta' Z(t) + W(t), t \in [0, 1] \tag{1}$$

where $W(t)$ is an unobservable standard Brownian motion, $Z(t) = (t, ..., t^p)'$ is a time polynomial vector and β is a parameter vector to be estimated.

The following two estimators of β are proposed:

$$\hat{\beta} = \left(\int_0^1 Z(t) Z(t)' dt \right)^{-1} \left(\int_0^1 Z(t) X(t) dt \right),$$

and

$$\tilde{\beta} = \left(\int_0^1 Z^{(1)}(t) Z^{(1)}(t)' dt \right)^{-1} \left(\int_0^1 Z^{(1)}(t) dX(t) \right),$$

where $Z^{(1)}$ is the vector of the first derivatives of Z.

Part A

1. Show that both $\hat{\beta}' Z(t)$ and $\tilde{\beta}' Z(t)$ are Hilbert projections in $L_2[0, 1]$. How do these projections differ?

2. Find the distributions of $\hat{\beta}$ and $\tilde{\beta}$ and compare them in the case where $p = 1$. What do you conclude?

Part B

Suppose the system generating $X(t)$ is

$$X(t) = \beta' Z(t) + J_c(t), \ \ t \in [0, 1] \tag{2}$$

where $J_c(t) = \int_0^t e^{(t-s)c} dW(s)$ for some known constant $c < 0$, is a linear diffusion, or Orstein-Uhlenbeck process.

1. Are $\hat{\beta}' Z(t)$ and $\tilde{\beta}' Z(t)$ still Hilbert projections?

2. Calculate the distributions of $\hat{\beta}$ and $\tilde{\beta}$ and compare them in the case where $p = 1$. What do you conclude?

3. Taking c to be known, can you suggest an unbiased linear estimator of β which has smaller variance than $\hat{\beta}$ and $\tilde{\beta}$? Does it correspond to another Hilbert projection?

2. Solution

Part A: Brownian Motion Case

Define the operator P as

$$PX(t) = \widehat{\beta}' Z(t) = \left(\int_0^1 X(s) Z(s)' ds \right) \left(\int_0^1 Z(s) Z(s)' ds \right)^{-1} Z(t),$$

then

$$\begin{aligned} P^2 X &= \left(\int_0^1 PX(s) Z(s)' ds \right) \left(\int_0^1 Z(s) Z(s)' ds \right)^{-1} Z(t) \\ &= \left[\left(\int_0^1 X Z' ds \right) \left(\int_0^1 Z Z' ds \right)^{-1} \int_0^1 Z Z' ds \right] \left(\int_0^1 Z Z' ds \right)^{-1} Z(t) \\ &= \left(\int_0^1 X(s) Z(s)' ds \right) \left(\int_0^1 Z(s) Z(s)' ds \right)^{-1} Z(t) = PX. \end{aligned}$$

In addition, it is easy to show that $(PX, Y) = (PY, X)$. Thus P is an orthogonal Hilbert projector onto the manifold spanned by $Z(t)$.

Define the operator Q as

$$QX = \widetilde{\beta}' Z = \left(\int_0^1 dX(s) Z^{(1)}(s)' \right) \left(\int_0^1 Z^{(1)}(s) Z^{(1)}(s)' ds \right)^{-1} Z(t).$$

Then Q is idempotent because

$$\begin{aligned} Q^2 X &= \left(\int_0^1 d[QX] Z^{(1)}(s)' \right) \left(\int_0^1 Z^{(1)}(s) Z^{(1)}(s)' ds \right)^{-1} Z(t) \\ &= \left(\int_0^1 dX(s) Z^{(1)}(s)' \right) \left(\int_0^1 Z^{(1)}(s) Z^{(1)}(s)' ds \right)^{-1} \\ &\quad \left(\int_0^1 dZ(s) Z^{(1)}(s)' \right) \left(\int_0^1 Z^{(1)}(s) Z^{(1)}(s)' ds \right)^{-1} Z(t) \\ &= \left(\int_0^1 dX(s) Z^{(1)}(s)' \right) \left(\int_0^1 Z^{(1)}(s) Z^{(1)}(s)' ds \right)^{-1} Z(t) \\ &= QX. \end{aligned}$$

But Q is not orthogonal since

$$(QX, Y) = \left(\int_0^1 dX(s) Z^{(1)}(s)' ds \right) \left(\int_0^1 Z^{(1)}(s) Z^{(1)}(s)' ds \right)^{-1} \left(\int_0^1 Z(s) Y(s) ds \right),$$

while

$$(X, QY) = \left(\int_0^1 dY(s)Z^{(1)}(s)'ds\right)\left(\int_0^1 Z^{(1)}(s)Z^{(1)}(s)'ds\right)^{-1}\left(\int_0^1 Z(s)X(s)ds\right)$$

and so $(QX, Y) \neq (X, QY)$ in general.

Therefore both PX and QX are Hilbert projections of $X(t)$ onto the space spanned by $Z(t)$. The difference lies in that PX is an orthogonal projection while QX is a non-orthogonal projection. We now investigate the statistical properties of these two projections.

First consider the distribution of $\widehat{\beta}$. Since

$$\widehat{\beta} - \beta = \left(\int_0^1 Z(s)Z(s)'ds\right)^{-1}\int_0^1 Z(s)W(s)ds,$$

is a linear functional of a Brownian motion, we know that $\widehat{\beta} - \beta \sim N(0, V_1)$, where

$$V_1 = \left(\int_0^1 Z(s)Z(s)'ds\right)^{-1}\left(\int_0^1\int_0^1 Z(t)(s \wedge t)Z(s)'dtds\right)\left(\int_0^1 Z(s)Z(s)'ds\right)^{-1}.$$

Next consider the distribution of $\widetilde{\beta}$. Since

$$\widetilde{\beta} - \beta = \left(\int_0^1 Z^{(1)}(s)Z^{(1)}(s)'ds\right)^{-1}\int_0^1 Z^{(1)}(s)dW(s),$$

is also a linear functional of a Brownian motion, we know immediately that $\widetilde{\beta} - \beta \sim N(0, V_2)$, where

$$V_2 = \left(\int_0^1 Z^{(1)}(s)Z^{(1)}(s)'ds\right)^{-1}.$$

We now compare the distributions of $\hat{\beta}$ and $\tilde{\beta}$ when $p = 1$. In this case $V_1 = \frac{6}{5}$ and $V_2 = \int_0^1 dr = 1$. Therefore $\widetilde{\beta}$ is a more efficient estimator than $\widehat{\beta}$ and the efficiency gain is 20%.

Remark

Note that $\hat{\beta}'Z(t)$ is the best $L_2[0,1]$ functional approximation to $\beta'Z(t)$ and it is delivered by an orthogonal projection of $X(t)$ onto the manifold spanned by $Z(t)$. On the other hand, $\tilde{\beta}'Z(t)$ is the better (more efficient) statistical estimator of $\beta'Z(t)$ and it is delivered by a non-orthogonal projection onto the same space.

Part B: The Diffusion Process Case

Both $\widehat{\beta}' Z(t)$ and $\widetilde{\beta}' Z(t)$ are still Hilbert projections. The proofs are essentially the same as before and omitted.

We proceed to compute the distributions of $\hat{\beta}$ and $\tilde{\beta}$. First

$$\widehat{\beta}-\beta=\left(\int_0^1 Z(s)Z(s)'ds\right)^{-1}\int_0^1 Z(t)J_c(t)dt \sim N(0,V_3),$$

where

$$\begin{aligned} V_3 &= \left(\int_0^1 Z(s)Z(s)'ds\right)^{-1} Var\left[\int_0^1 Z(t)J_c(t)dt\right]\left(\int_0^1 Z(s)Z(s)'ds\right)^{-1} \\ &= \left(\int_0^1 ZZ'ds\right)^{-1}\left\{\int_0^1\left(\int_0^1 Z(r)\Lambda_{r,s}Z(s)'ds\right)dr\right\}\left(\int_0^1 ZZ'ds\right)^{-1}. \end{aligned}$$

and $\Lambda_{r,s}=EJ_c(r)J_c(s)=\left(e^{(r+s)c}-e^{|r-s|c}\right)/(2c)$ is the covariance kernel of the O-U process.

Next, the error process $J_c(t)=\int_0^t e^{(t-s)c}dW(s)$ in (2) satisfies the linear differential equation

$$dJ_c(t)=cJ_c(t)dt+dW(t),$$

so we can write $\widetilde{\beta}-\beta$ as

$$\widetilde{\beta}-\beta=\left(\int_0^1 Z^{(1)}(s)Z^{(1)}(s)'ds\right)^{-1}\left(Z^{(1)}(1)J_c(1)-\int_0^1 Z^{(2)}(s)J_c(s)ds\right)\sim N(0,V_4),$$

where the equality follows from integration by parts and

$$\begin{aligned} V_4 &= \left(\int_0^1 Z^{(1)}Z^{(1)\prime}\right)^{-1}\left(\int_0^1\int_0^1 Z^{(2)}(r)\Lambda_{r,s}Z^{(2)}(s)'dsdr\right. \\ &\quad \left.+Z^{(1)}(1)\Lambda_{1,1}Z^{(1)}(1)'-2\int_0^1 Z^{(1)}(1)\Lambda_{1,s}Z^{(2)}(s)'ds\right)\left(\int_0^1 Z^{(1)}Z^{(1)\prime}\right)^{-1}. \end{aligned}$$

When $p=1$, some algebra gives $V_3=\frac{3}{2c^5}[2c^3+3c^2-3+3(c-1)^2e^{2c}]$ and $V_4=\Lambda_{1,1}=\frac{e^{2c}-1}{2c}$. So $V_3/V_4=3[2c^3+3c^2-3+3(c-1)^2e^{2c}]/[c^4(e^{2c}-1)]$. Figure 1 graphs the relative efficiency of $\widetilde{\beta}$ and $\widehat{\beta}$, *i.e.* V_3/V_4 against c. It reveals that $\widehat{\beta}$ is more efficient than $\widetilde{\beta}$ for small values of c. As c approaches zero, $\widetilde{\beta}$ becomes more efficient, which is consistent with our results in Part A. As c approaches zero from below, the linear diffusion J_c behaves like the underlying Brownian motion and the estimator delivered by the non-orthogonal projection is more efficient. The cutoff value of c is approximately -3.80. When $c<-3.80, V_3<V_4$; When $-3.80\le c<0$, $V_3\ge V_4$.

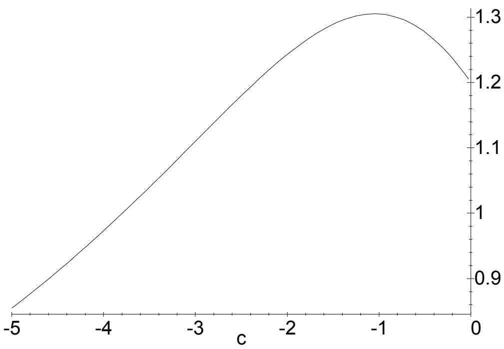

Figure 1

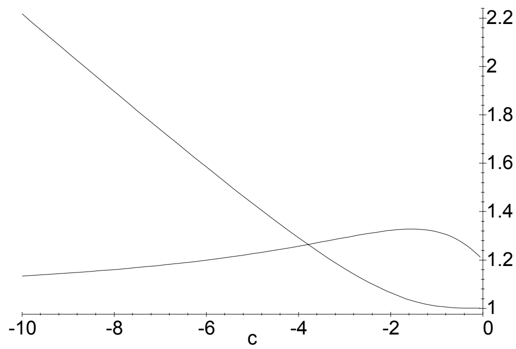

Figure 2

A More Efficient Estimator

We now propose another unbiased estimator $\overline{\beta}$, which is more efficient than both $\hat{\beta}$ and $\tilde{\beta}$. Define $Z_c(t) = Z^{(1)}(t) - cZ(t)$ and

$$\overline{\beta} = \left(\int_0^1 Z_c(t)Z_c(t)'dt\right)^{-1}\left(\int_0^1 Z_c(t)dX\,(t) - c\int_0^1 Z_c(t)X\,(t)\,dt\right). \qquad (3)$$

Recall that $dJ_c(t) = cJ_c(t)dt + dW\,(t)$, so we may linearly transform the model (2) to

$$dX(t) - cX\,(t)\,dt = \beta'\left[Z^{(1)}(t) - cZ\,(t)\right]dt + dW\,(t)\,,\ \ t \in [0,1], \qquad (4)$$

in which the innovations $dW\,(t)$ are independent. The estimator $\overline{\beta}$ in (3) is the 'least squares' estimator of β in the system (4). It follows from (3) and (4) that

$$\overline{\beta} - \beta = \left(\int_0^1 Z_c(t)Z_c(t)'dt\right)^{-1}\left(\int_0^1 Z_c(t)dW\,(t)\right) \sim N(0, V_5),$$

where

$$V_5 = \left(\int_0^1 Z_c(t)Z_c(t)'dt\right)^{-1}.$$

When $p = 1$, $V_5 = \left(\int_0^1 (ct-1)^2 dt\right)^{-1} = \frac{3}{c^2 - 3c + 3}$. Figure 2 graphs V_3/V_5 and V_4/V_5 against c. It shows that $\overline{\beta}$ is indeed more efficient than both $\widehat{\beta}$ and $\widetilde{\beta}$ for all values of $c < 0$.

Define the operator M as

$$MX = \left(\int_0^1 dX\,(t)\,Z_c(t)' - c\int_0^1 X\,(t)\,Z_c(t)'dt\right)\left(\int_0^1 Z_c(t)Z_c(t)'dt\right)^{-1} Z(t). \qquad (5)$$

Then, since $Z\,(t)$ is continuously differentiable

$$\int_0^1 dZ\,(t)\,Z_c(t)' - c\int_0^1 Z\,(t)\,Z_c(t)'dt = \int_0^1 Z_c(t)Z_c(t)'dt,$$

we have

$$MZ = \left(\int_0^1 dZ\,(t)\, Z_c(t)' - c\int_0^1 Z\,(t)\, Z_c(t)' dt\right)\left(\int_0^1 Z_c(t) Z_c(t)' dt\right)^{-1} Z(t) = Z\,(t)$$

and by virtue of the definition (5)

$$\begin{aligned} & M^2 X = MMX \\ = & \left(\int_0^1 dX\,(t)\, Z_c(t)' - c\int_0^1 X\,(t)\, Z_c(t)' dt\right)\left(\int_0^1 Z_c(t) Z_c(t)' dt\right)^{-1} MZ(t) \\ = & \left(\int_0^1 dX\,(t)\, Z_c(t)' - c\int_0^1 X\,(t)\, Z_c(t)' dt\right)\left(\int_0^1 Z_c(t) Z_c(t)' dt\right)^{-1} Z(t) \\ = & MX. \end{aligned}$$

So M is a Hilbert projection. In addition, it is easy to see that $(MX, Y) \neq (X, MY)$ in general. Therefore, $\overline{\beta}' Z(t) = MX$ is a non-orthogonal Hilbert projection. However it delivers the minimum variance unbiased trend estimator in this case.

Remarks

(a) The transformed system (4) is the continuous time version of the traditional Cochrane-Orcutt transformation of a discrete time regression model to remove autoregressive errors.

(b) The results obtained here clearly apply for any differentiable deterministic sequence $Z\,(t)$ for which $\int_0^1 ZZ' dt$, $\int_0^1 Z^{(1)} Z^{(1)}$, and $\int_0^1 Z_c Z_c' dt$ are positive definite.