Program in Applied Mathematics

Data-driven Model Reduction by Wiener Projection

Jared McBride

Applied Mathematics
University of Arizona

Mar 6, 2020

THE UNIVERSITY OF ARIZONA

Outline

The Problem

Many important models today contemplate

- large number of degrees of freedom
- across many orders of magnitude in space and time, without sharp scale separation.

Examples: Power flow on large grids, neural activity in the brain, weather forecasting, etc.

Some tasks require repeated model runs such as for

- Uncertainty quantification
- Optimization and control

Commonly, only a relatively small number of variables are of direct interest or even observable.

Goal and Difficulties

The **goal** is then to find reduced order models, which include only the variables of interest (resolved variables), capable of finite time forecasting as well as reproducing long-time statistics like correlation functions and marginals of stationary distributions, at lower computational costs.

How can the effect of unresolved variables be approximated by using the resolved variables and stochastic terms.

Unlike under situations with sharp scale separation,

- memory (marginals of Markov process may not be Markov) and
- noise effects

must be accounted for in many applications.

The object of interest are signals, realizations from a stochastic process.

We will focus on discrete time.

Some terms

- A (linear) filter is in this talk a sequence of deterministic elements $h_{n,k}$, $n, k > -\infty$ that operates on a signal by way of convolution.

$$\hat{\mathbf{y}}_n = (\mathbf{y} \star h_{n,\cdot})_n = \sum_{k=-\infty}^{\infty} h_{n,k} \cdot \mathbf{y}_{n-k}$$

- A linear filter is time-invariant if $h_{n,k}$ has no dependence on n (used above). So, $h_{n,k} = h_k$
- A linear time-invariant filter is causal if $h_k = 0$ for $k < 0$.

A stochastic process $\mathbf{x}_t$ is (wide sense) stationary if

- $m(t) = \mathbb{E}\mathbf{x}_t$ = constant and
- $R_{\mathbf{x}}(t,s) = \mathbb{E}\Big[(\mathbf{x}_t - \mathbb{E}\mathbf{x}_t)(\mathbf{x}_s - \mathbb{E}\mathbf{x}_s)^*\Big] = R_{\mathbf{x}}(t-s)$.

The autocovariance of a stationary process $\mathbf{x}_t$ is the function

$$R_{\mathbf{x}}(t-s) = \mathbb{E}\Big[(\mathbf{x}_t - \mathbb{E}\mathbf{x}_t)(\mathbf{x}_s - \mathbb{E}\mathbf{x}_s)^*\Big].$$

Observe if $\mathbf{x}_t$ is vector-valued then the autocovariance is matrix valued.

- $R_{\mathbf{x}}(0) \geq 0$ (positive definite)
- $R_{\mathbf{x}}(t) = R_{\mathbf{x}}(-t)^*$

Two stochastic processes $\mathbf{x}_t$ and $\mathbf{y}_t$ are jointly stationary if the process $\mathbf{u}_t = (\mathbf{x}_t, \mathbf{y}_t)$ is stationary. This means $R_{\mathbf{u}}(t-s)$ has a block representation

$$R_{\mathbf{u}}(t-s) = \begin{pmatrix} \mathbb{E}\Big[(\mathbf{x}_t - \mathbb{E}\mathbf{x}_t)(\mathbf{x}_s - \mathbb{E}\mathbf{x}_s)^*\Big] & \mathbb{E}\Big[(\mathbf{x}_t - \mathbb{E}\mathbf{x}_t)(\mathbf{y}_s - \mathbb{E}\mathbf{y}_s)^*\Big] \\ \mathbb{E}\Big[(\mathbf{y}_t - \mathbb{E}\mathbf{y}_t)(\mathbf{x}_s - \mathbb{E}\mathbf{x}_s)^*\Big] & \mathbb{E}\Big[(\mathbf{y}_t - \mathbb{E}\mathbf{y}_t)(\mathbf{y}_s - \mathbb{E}\mathbf{y}_s)^*\Big] \end{pmatrix}$$
$$=: \begin{pmatrix} R_{\mathbf{x}}(t-s) & R_{\mathbf{xy}}(t-s) \\ R_{\mathbf{yx}}(t-s) & R_{\mathbf{y}}(t-s) \end{pmatrix}$$

The cross-covariance of two jointly stationary processes $\mathbf{x}_t$, $\mathbf{y}_t$ is the function

$$R_{\mathbf{yx}}(t-s) = \mathbb{E}\Big[(\mathbf{y}_t - \mathbb{E}\mathbf{y}_t)(\mathbf{x}_s - \mathbb{E}\mathbf{x}_s)^*\Big].$$

Observe that

$$R_{\mathbf{yx}}(t-s) = R^*_{\mathbf{yx}}(s-t)$$

The spectral density of a stationary discrete-time process $\mathbf{x}_n$ is the function

$$\hat{S}_{\mathbf{x}}(\omega) = \sum_{n=-\infty}^{\infty} R_{\mathbf{x}}(n) e^{-i\omega n}$$

Since $R_{\mathbf{x}}(n) = R_{\mathbf{x}}(-n)^*$, $\hat{S}_{\mathbf{x}}(\omega) \in \mathbb{R}$ for all $\omega \in \mathbb{R}$,
in fact $\hat{S}_{\mathbf{x}}(\omega) \geq 0$ for all $\omega \in \mathbb{R}$.

The z-spectrum of a stationary discrete-time process $\mathbf{x}_t$ is the function

$$S_{\mathbf{x}}(z) = \sum_{n=-\infty}^{\infty} R_{\mathbf{x}}(n) z^{-n} \quad \left(= \mathcal{Z}\{R_{\mathbf{x}}(n)\} \right)$$

$S_{\mathbf{x}}(z)$ is positive semi-definite on the unit circle.

The z-cross-spectrum of two jointly stationary discrete-time processes $\mathbf{y}_n$ and $\mathbf{x}_n$ is the function

$$S_{\mathbf{yx}}(z) = \sum_{n=-\infty}^{\infty} R_{\mathbf{yx}}(n) z^{-n} = \left[S_{\mathbf{xy}}\left(\frac{1}{z^*}\right)\right]^* = S^*_{\mathbf{xy}}(z^{-*})$$

For a time series $x = (x_n, n \geq 0)$ we define the z-transform of x by

$$X(z) = \sum_{n=0}^{\infty} x_n z^{-n} = \mathcal{Z}\{x_n\}$$

Here are some properties:

$$\mathcal{Z}\{x_{n+1}\} = \sum_{n=0}^{\infty} x_{n+1} z^{-n} = \sum_{n=1}^{\infty} x_n z^{-n+1} = z\left(\sum_{n=0}^{\infty} x_n z^{-n} - x_0\right) = z\left(X(z) - x_0\right)$$

$$\mathcal{Z}\{x_{n-k}\} = \sum_{n=0}^{\infty} x_{n-k} z^{-n} = \sum_{n=-k}^{\infty} x_n z^{-n-k} = z^{-k}\left(\sum_{n=0}^{\infty} x_n z^{-n}\right) = z^{-k} X(z)$$

(more useful properties)

$$x_n = \mathcal{Z}^{-1}\{X(z)\} = \frac{1}{2\pi i}\int_C X(z) z^{n-1}\, dz$$

$$\begin{aligned}
\mathcal{Z}\{(x \star y)_n\} = \mathcal{Z}\left\{\sum_{k=0}^{\infty} x_{n-k}\cdot y_k\right\} &= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{\infty} x_{n-k}\cdot y_k\right) z^{-n} \\
&= \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} x_{n-k} z^{-n}\cdot y_k \\
&= \sum_{k=0}\left(\sum_{n=0}^{\infty} x_n z^{-n}\right)\cdot y_k z^{-k} \\
&= \left(\sum_{n=0}^{\infty} x_n z^{-n}\right)\cdot\left(\sum_{k=0} y_k z^{-k}\right) \\
&= X(z)\cdot Y(z).
\end{aligned}$$

(more useful properties)

$$x_n = \mathcal{Z}^{-1}\{X(z)\} = \frac{1}{2\pi i}\int_C X(z) z^{n-1}\, dz$$

$$\begin{aligned}
\mathcal{Z}\{(x \star y)_n\} = \mathcal{Z}\left\{\sum_{k=0}^{\infty} x_{n-k}\cdot y_k\right\} &= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{\infty} x_{n-k}\cdot y_k\right) z^{-n} \\
&= \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} x_{n-k} z^{-n}\cdot y_k \\
&= \sum_{k=0}\left(\sum_{n=0}^{\infty} x_n z^{-n}\right)\cdot y_k z^{-k} \\
&= \left(\sum_{n=0}^{\infty} x_n z^{-n}\right)\cdot\left(\sum_{k=0} y_k z^{-k}\right) \\
&= X(z)\cdot Y(z).
\end{aligned}$$

Keep in mind that this is true for matrix and/or vector-valued time series in which the multiplication makes sense.

If $\mathbf{x}$ and $\mathbf{y}$ are two stationary processes, then

$$S_{\mathbf{x}+\mathbf{y}}(z) = S_{\mathbf{x}}(z) + S_{\mathbf{xy}}(z) + S_{\mathbf{yx}}(z) + S_{\mathbf{y}}(z)$$

If $\mathbf{x}$ and $\mathbf{y}$ are uncorrelated $S_{\mathbf{yx}}(z) = 0$ and

$$S_{\mathbf{x}+\mathbf{y}}(z) = S_{\mathbf{x}}(z) + S_{\mathbf{y}}(z)$$

If $\mathbf{u}_n \sim N(0, \sigma_{\mathbf{u}})$ then

$$R_{\mathbf{u}}(n) = \sigma_{\mathbf{u}}^2 \delta(n)$$
$$S_{\mathbf{u}}(z) = \sigma_{\mathbf{u}}^2$$

Program in Applied Mathematics

Suppose $\mathbf{x}_n$ is obtained from a stationary signal $\mathbf{y}_n$ by passing it through a linear time-invariant filter w_n, so $\boxed{\mathbf{x}_n = (\mathbf{y} \star w)_n}$.
Then, by the convolution theorem,

$$S_{\mathbf{yx}}(z) = S_{\mathbf{y}}(z) W^*(z^{-*})$$

and

$$S_{\mathbf{x}}(z) = W(z) S_{\mathbf{y}}(z) W^*(z^{-*})$$

where

$$W(z) = \mathcal{Z}\{w_n\} = \sum_{n=-\infty}^{\infty} w_n z^{-n}$$

and is called the transfer function of the filter w.

Wiener's Matrix Spectral Factorization Theorem

If $S : \mathbb{C} \to \mathbb{C}^{d \times d}$, satisfies,

- $S \in L^1(\partial \mathbb{D})$,
- $\log \det S \in L^1(\partial \mathbb{D})$, and
- $S(z) > 0$ (positive definite) for (almost all) $z \in \partial \mathbb{D}$.

Then there exists matrix functions $S^+(z)$ and $S^-(z)$, such that $S^-(z) = S^{+*}(z^{-*})$ and

$$S(z) = S^+(z) S^-(z) \qquad \text{for } z \in \partial \mathbb{D}.$$

Furthermore, S^+ is is an outer analytic matrix function from the Hardy space H_2.

More useful version of Spectral Factorization Theorem

If $\mathbf{y}$ is a mean zero, stationary, discrete time stochastic d-vector-valued process that admits a rational z-spectrum $S_{\mathbf{y}}$ analytic on some annulus containing the unit circle, and

$$S_{\mathbf{y}} > 0 \qquad \text{everywhere on } \partial\mathbb{D}.$$

Then there exists matrix functions $S^{+}(z)$ and $S^{-}(z)$, such

- $S^{+}(z)$ is a $d \times d$ rational matrix function that is analytic on and inside the unit circle,
- $S^{+-1}(z)$ is analytic on and inside the unit circle.
- $S^{-}(z) = S^{+*}(z^{-*})$ and
- $S(z) = S^{+}(z)S^{-}(z)$.

Program in Applied Mathematics

Most Numerical algorithms assume $S(z)$ is rational and has the form of a Laurent Polynomial meaning it may be written as

$$S(z) = \sum_{n=-m}^{m} c_n z^{-n} \qquad \text{with } c_n = c^*_{-n}.$$

If this is assumed it may be shown that

$$S^+(z) = \sum_{n=1} L_n z^n \qquad \text{and} \qquad S^-(z) = \sum_{n=1} L^*_n z^{-n}$$

(this is what we assume here) Algorithms that use Toeplitz matrices.

- Bauer
- Schur
- Levinson-Durbin

Algorithms that use State Space formulations.

- Riccati Equation
- Kalman Filter
- Chadrasekhar-Kailath-Morf-Sidhu (CKMS)

Recently Analgroithm that imposes no more than the general theorem

2318 IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 57, NO. 4, APRIL 2011

A New Method of Matrix Spectral Factorization

Gigla Janashia, Edem Lagvilava, and Lasha Ephremidze

***Abstract*—A new algorithm of matrix spectral factorization is proposed which can be applied to compute an approximate spectral factor of any positive definite matrix function which satisfies the Paley-Wiener condition.** [1]

***Index Terms*—Algorithms, matrix decomposition, spectral factorization.**

I. Introduction

SPECTRAL factorization plays a prominent role in a wide range of fields in Communications, System Theory, Control Engineering and so on. In the scalar case arising for single input and single output systems, the factorization problem is relatively easy and several classical algorithms exist to tackle it (see the survey paper [1]) together with the reliable information on their software implementations [2]. There are also some recent claims as to their improvement [3]. Matrix spectral factorization which arises for multidimensional systems is essentially this process the decisive role is played by unitary matrix functions of certain structure (see Theorem 1), which eliminates many technical difficulties connected with computation. The explicit construction of such matrices given in the proof of Theorem 1 is an essential component of the algorithm. Recently, a close relationship of these unitary matrix functions with compactly supported wavelets has been discovered, which makes it possible to construct compact wavelets in a fast and reliable way and to completely parameterize them [8].

Preliminary numerical simulations confirm the potential of the proposed algorithm (see Section VI).

The paper is organized as follows. In the next section, an exact mathematical formulation of the problem is given. In Section III, the notation used throughout the paper is introduced. In Section IV, we provide a theoretical background of the proposed method. In Section V, the computational procedures of the new matrix spectral factorization algorithm are described, and some illustrating examples of numerical

Given two stationary processes $\mathbf{x}_n$,$\mathbf{y}_n$ The Wiener Filter computes a linear least square estimate $\hat{\mathbf{y}}_n$ of a process $\mathbf{y}_n$ given $\mathbf{x}_n$, for this reason

- $\mathbf{y}_n$ is called the signal,
- $\mathbf{x}_n$ are called the predictors.

This means we seek an h such that

$$\mathbb{E}\|\mathbf{y}_n - (\mathbf{x} \star h)_n\|^2 = \text{minimum}$$

In our case we want to require h_n to be

- causal (meaning $h_n = 0$ for $n < 0$)
- rapid decay (so that efficiency is gained)

We assume we have it, but that h was not assumed to be causal. Then

$$0 = \mathbb{E}\left[(\mathbf{y}_n - \hat{\mathbf{y}}_n)(\mathbf{x}_m)\right] = \mathbb{E}\left[(\mathbf{y}_n - (h \star \mathbf{x})_n)(x_m)^*\right]$$

This implies

$$\mathbb{E}\mathbf{y}_n\mathbf{x}_m^* = \mathbb{E}\sum_{k=-\infty}^{\infty} h_k \mathbf{x}_{n-k}\mathbf{x}_m^* = \sum_{k=-\infty}^{\infty} h_k \mathbb{E}\mathbf{x}_{n-k}\mathbf{x}_m^*$$

or rather (with relabeling $n - m \mapsto n$)

$$R_{\mathbf{yx}}(n) = \sum_{k=-\infty}^{\infty} h_k R_{\mathbf{x}}(n-k)$$

The form of RHS suggest use of the z-transform.

Applying the z-transform to both sides gives

$$S_{\mathbf{yx}}(z) = H(z)S_{\mathbf{x}}(z)$$

where

$$H(z) = \mathcal{Z}\{h_n\} = \sum_{n=-\infty}^{\infty} h_n z^{-n}$$

So,

$$H(z) = S_{\mathbf{yx}}(z)S_{\mathbf{x}}^{-1}(z)$$

we then apply the inverse z-transform to recover h

$$h_n = \frac{1}{2\pi i}\int_C S_{\mathbf{yx}}(z)S_{\mathbf{x}}^{-1}(z)z^{n-1}\,dz$$

If we require that h is causal this is more difficult. Then

$$0 = \mathbb{E}\left[(\mathbf{y}_n - \hat{\mathbf{y}}_n)(\mathbf{x}_m)\right] = \mathbb{E}\left[(\mathbf{y}_n - (h \star \mathbf{x})_n)(x_m)^*\right] \qquad \text{only for } m \leq n$$

This implies

$$\mathbb{E}\mathbf{y}_n\mathbf{x}_m^* = \mathbb{E}\sum_{k=-\infty}^{\infty} h_k \mathbf{x}_{n-k}\mathbf{x}_m^* = \sum_{k=-\infty}^{\infty} h_k \mathbb{E}\mathbf{x}_{n-k}\mathbf{x}_m^* \qquad \text{only for } m \leq n$$

or rather (with relabeling $n - m \mapsto n$)

$$R_{\mathbf{yx}}(n) = \sum_{k=-\infty}^{\infty} h_k R_{\mathbf{x}}(n-k) \qquad \text{only for } n \geq 0$$

The form of RHS suggest use of the z-transform. But we can't!

However observe that for

$$g_n = R_{\mathbf{yx}}(n) - \sum_{k=-\infty}^{\infty} h_k R_{\mathbf{x}}(n-k)$$

g is strictly anti-casual since $g_n = 0$ when $n \geq 0$. Now apply the z-transform to both sides. We get

$$G(z) = S_{\mathbf{yx}}(z) - H(z)S_{\mathbf{x}}(z)$$

Now apply the spectral factorization to $S_{\mathbf{x}}(z)$ And proceed as follows

$$G(z) = S_{\mathbf{yx}}(z) - H(z)S_{\mathbf{x}}^{-}(z)S_{\mathbf{x}}^{+}(z)$$

and observe when we apply the inverse

$$\underbrace{G(z){S^{+}_{\mathbf{x}}}^{-1}(z)}_{strictly\ anti-causal} = \underbrace{S_{\mathbf{yx}}(z){S^{+}_{\mathbf{x}}}^{-1}(z)}_{mixed} - \underbrace{H(z)S_{\mathbf{x}}^{-}(z)}_{causal}$$

And so

$$H(z) = \left\{ S_{\mathbf{yx}}(z) {S^{+}_{\mathbf{x}}}^{-1}(z) \right\}_{+} {S^{-}_{\mathbf{x}}}^{-1}(z)$$

This Approach

Given a full model

$$X_n = F(X_n)$$

with resolved variables collected in x_n, select functions $\psi^{(i)}(x)$ (informed by model) on reduced state variables.

$$\psi(x) = \left(\psi^{(0)}(x)\middle|\psi^{(1)}(x)\middle|\cdots\middle|\psi^{(\nu)}\right) \qquad \text{and} \qquad \psi_k = \psi(x_k).$$

The reduced model we purpose is of the form,

$$\boxed{x_{n+1} = \sum_{k=0}^{\infty} \psi_k \cdot h_{n-k} + \xi_{n+1}}$$

We use the data to infer h_k and ξ_n (described briefly below).

Dr. Kevin Lin and Dr. Fei Lu solves in time domain, using an iterative optimization algorithm.

This study investigates computing the Wiener filter by spectral methods (that is, employing information like the power spectra S_{yx} and S_x). This is a direct method requiring no iterative optimization

Advantages:

- Quicker
- more accurate (?)

Example 1: MA(1) Signal, Additive WN

Let $\mathbf{y} \sim \text{MA}(1)$ with $r \in \mathbb{R}$ have the form

$$\mathbf{y}_n = \mathbf{u}_n - r\mathbf{u}_{n-1} \qquad \text{for } n > -\infty$$

$\mathbf{u}_n \sim N(0,1)$ i.i.d. And let y_n be a realization, this will be the signal. The observations are

$$x_n = y_n + v_n, \qquad \text{for } n > -\infty.$$

we assume $\mathbf{v}_n \sim N(0, \sigma_{\mathbf{v}})$ i.i.d. and are uncorrelated with $\mathbf{y}$.

To compute the Wiener filter we need $S_{\mathbf{yx}}$, $S_{\mathbf{x}}^{+}$, and $S_{\mathbf{x}}^{-}$. First, observe

$$S_{\mathbf{yx}} = S_{\mathbf{y}} + S_{\mathbf{yv}} = S_{\mathbf{y}},$$

and

$$S_{\mathbf{x}} = S_{\mathbf{y+v,y+v}} = S_{\mathbf{y}} + S_{\mathbf{yv}} + S_{\mathbf{vy}} + S_{\mathbf{v}} = S_{\mathbf{y}} + \sigma_{\mathbf{v}}^2.$$

Example 1: MA(1) Signal, Additive WN

Program in Applied Mathematics

(Just this once we compute it for clarity)

$$
\begin{aligned}
S_{\mathbf{y}} &= \sum_{k=-\infty}^{\infty} \mathbb{E}[(u_{n+k} - ru_{n+k-1})(u_n - ru_{n-1})^*]z^{-k} \\
&= \sum_{k=-\infty}^{\infty} \mathbb{E}[u_{n+k}u_n^* - ru_{n+k-1}u_n^* - ru_{n+k}u_{n-1}^* + r^2 u_{n+k-1}u_{n-1}^*]z^{-k} \\
&= \sum_{k=-\infty}^{\infty} \big(\delta(k) - r\delta(k-1) - r\delta(k+1) + r^2\delta(k)\big)z^{-k} \\
&= 1 + r^2 - rz^{-1} - rz \qquad \big(\quad = (1-rz)(1-rz^{-1}) \quad \big).
\end{aligned}
$$

So, $S_{\mathbf{yx}}(z) = (1-rz)(1-rz^{-1})$ and

$$
\begin{aligned}
S_{\mathbf{x}}(z) &= 1 + r^2 + \sigma_{\mathbf{v}}^2 - rz^{-1} - rz \\
&= \frac{r}{\rho}(1-\rho z^{-1})(1-\rho z)
\end{aligned}
$$

For a suitably chosen ρ, $|\rho| < 1$.

For $S_{\mathbf{x}}^{+}$ and $S_{\mathbf{x}}^{-}$ we then get

$$S_{\mathbf{x}}^{+}(z) = \sqrt{\frac{r}{\rho}}(1 - \rho z) \qquad \text{and} \qquad S_{\mathbf{x}}^{-}(z) = \sqrt{\frac{r}{\rho}}(1 - \rho z^{-1}).$$

Putting this together we get

$$\begin{aligned}\frac{S_{\mathbf{yx}}(z)}{S_{\mathbf{x}}^{+}(z)} &= \frac{1 + r^2 - rz^{-1} - rz}{\sqrt{\frac{r}{\rho}}(1 - \rho z)} \\ &= \sqrt{\frac{\rho}{r}}(1 + r^2 - rz^{-1} - rz)\sum_{n=0}^{\infty}(\rho z)^n \\ &= -\sqrt{\rho r} z^{-1} + \sqrt{\frac{\rho}{r}}(1 + r^2) - \rho\sqrt{\rho r} + \sum_{n=1}^{\infty}\xi_n z^n.\end{aligned}$$

Which means

$$\left\{\frac{S_{\mathbf{yx}}(z)}{S_{\mathbf{x}}^{+}(z)}\right\}_{+} = -\sqrt{\rho r} z^{-1} + \sqrt{\frac{\rho}{r}}(1 + r^2) - \rho\sqrt{\rho r}.$$

$$H(z) = \frac{1}{S_{\mathbf{x}}^{-}(z)} \left\{ \frac{S_{\mathbf{yx}(z)}}{S_{\mathbf{x}}^{+}(z)} \right\}_{+} = \frac{1}{\sqrt{\frac{r}{\rho}}(1-\rho z^{-1})} \left(\sqrt{\rho r} z^{-1} + \sqrt{\frac{\rho}{r}}(1+r^2) - \rho\sqrt{\rho r} \right)$$

$$= \left[\frac{1+r^2}{r}\rho - \rho^2 \right] + \sum_{n=1}^{\infty} \rho^n \left[\frac{1+r^2}{r}\rho - \rho^2 - 1 \right] z^{-n}$$

If we let $d = \frac{1+r^2}{r}\rho - \rho^2$, then

$$H(z) = d + \sum_{n=1}^{\infty} \rho^n (d-1) z^{-n}$$

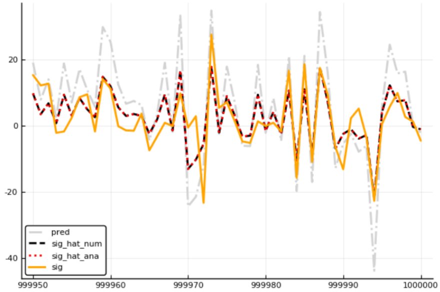
And the causal filter $h = (h_n, n > -\infty)$ is

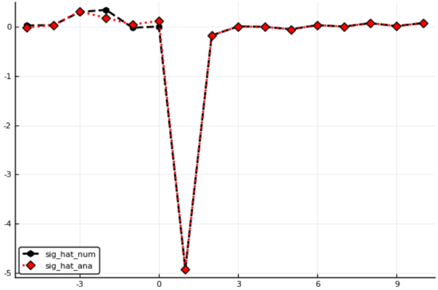
$$\begin{aligned} h_n &= 0 && \text{if } n < 0 \\ h_n &= d && \text{if } n = 0 \\ h_n &= \rho^n (d-1) && \text{if } n > 0 \end{aligned}$$

Example 1: MA(1) Signal, Additive WN

Program in Applied Mathematics

Here is a run, with $r = 10$, $\sigma_{\mathbf{v}} = 10$. The trajectory has 10^6 steps after discarding 10^3 steps.





Left: A window of the time series for the signal (orange), the predictors (light gray), the estimated signal using the analytic and numerical Wiener filter (red, black).

Right: The covariance between errors (red from analytic filter, black from numerical) and predictors (observations).

For the signal in this example we use the MA(2) process,

$$y_n = u_n + r_1 u_{n-1} + r_2 u_{n-2}, \qquad \text{for } n > -\infty$$

where $r_1, r_2 \in \mathbb{R}$. The observations will again simply be the signal with an additive white noise,

$$x_n = y_n + v_n, \qquad \text{for } n > -\infty.$$

We assume that $\mathbf{v} = (\mathbf{v}_n, n > -\infty)$ is uncorrelated with $\mathbf{y}$.

Skipping ahead we have

$$S_{\mathbf{yx}} = 1 + r_1^2 + r_2^2 + (r_1 + r_1 r_2)(z + z^{-1}) + r_2(z^2 + z^{-2}),$$

$$S_{\mathbf{x}}^{+}(z) = \sqrt{\frac{r_2}{\rho_1\rho_2}}(1-\rho_1 z)(1-\rho_2 z) \quad \text{and} \quad S_{\mathbf{x}}^{-}(z) = \sqrt{\frac{r_2}{\rho_1\rho_2}}(1-\rho_1 z^{-1})(1-\rho_2 z^{-1})$$

Example 2: MA(2) Signal, Additive WN

Program in Applied Mathematics

$$H(z) = \frac{1}{S_{\mathbf{x}}^{-}(z)} \left\{ \frac{S_{\mathbf{yx}}(z)}{S_{\mathbf{x}}^{+}(z)} \right\}_{+} = \frac{\sqrt{\frac{r_2}{\rho_1\rho_2}}(\alpha_2 z^{-2} + \alpha_1 z^{-1} + \alpha_0)}{\sqrt{\frac{r_2}{\rho_1\rho_2}}(1 - \rho_1 z^{-1})(1 - \rho_2 z^{-1})}$$

$$= (\alpha_2 z^{-2} + \alpha_1 z^{-1} + \alpha_0) \left(\sum_{n=0}^{\infty} \rho_1^n z^{-n} \right) \left(\sum_{m=0}^{\infty} \rho_2^m z^{-m} \right)$$

$$= \alpha_0 \alpha(0) + [\alpha_0 \alpha(1) + \alpha_1 \alpha(0)] z^{-1}$$

$$+ \sum_{n=2}^{\infty} [\alpha_0 \alpha(n) + \alpha_1 \alpha(n-1) + \alpha_2 \alpha(n-2)] z^{-n}$$

where $\alpha(n) = \sum_{k=0}^{n} \rho_1^{n-k} \rho_2^k$ and

$$\alpha_0 = \frac{\rho_1\rho_2}{r_2} \left[r_2(\rho_1^2 + \rho_2^2 + \rho_1\rho_2) + (r_1 + r_1 r_2)(\rho_1 + \rho_2) + 1 + r_1^2 + r_2^2 \right],$$

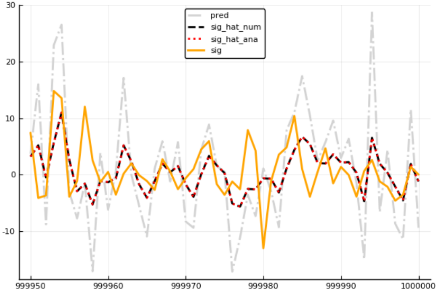
$$\alpha_1 = \frac{\rho_1\rho_2}{r_2} \left[r_2(\rho_1 + \rho_2) + r_1 + r_1 r_2 \right], \qquad \alpha_2 = \frac{\rho_1\rho_2}{r_2} r_2.$$

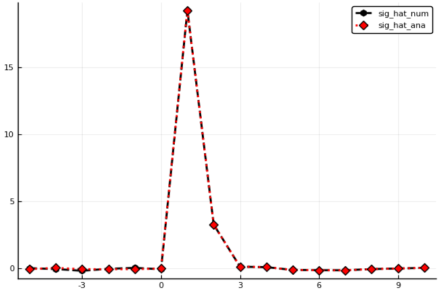
And the causal filter $h = (h_n, n > -\infty)$ is

$$
\begin{aligned}
h_n &= 0 && \text{if } n < 0 \\
h_n &= \alpha_0 \alpha(0) && \text{if } n = 0 \\
h_n &= \alpha_0 \alpha(1) + \alpha_1 \alpha(0) && \text{if } n = 1 \\
h_n &= \alpha_0 \alpha(n) + \alpha_1 \alpha(n-1) + \alpha_2 \alpha(n-2) && \text{if } n \geq 2
\end{aligned}
$$

Example 2: MA(2) Signal, Additive WN

Here is a run, with $r = 10$, $\sigma_{\mathbf{v}} = 10$. The trajectory has 10^6 steps after discarding 10^3 steps.





Left: A window of the time series for the signal (orange), the predictors (light gray), the estimated signal using the analytic and numerical Wiener filter (red, black).

Right: The covariance between errors (red from analytic filter, black from numerical) and predictors (observations).

Example 3: AR(2) Signal, Additive WN

Program in Applied Mathematics

Let us consider the stationary autoregressive process of order 2,

$$y_n = (r_1 + r_2)y_{n-1} - r_1 r_2 y_{n-2} + u_n, \qquad \text{for } n > -\infty$$

for $r_1, r_2 \in \{z : |z| < 1\}$.

Skipping way ahead we have

$$S_{\mathbf{yx}} = \frac{1}{(1 - r_1 z^{-1})(1 - r_2 z^{-1})(1 - r_1^* z)(1 - r_2^* z)},$$

$$S_{\mathbf{x}}^{+}(z) = \sqrt{\sigma_{\mathbf{v}}^2 \frac{r_1^* r_2^*}{\rho_1^* \rho_2^*}} \frac{(1 - \rho_1^* z)(1 - \rho_2^* z)}{(1 - r_1^* z)(1 - r_2^* z)},$$

$$S_{\mathbf{x}}^{-}(z) = \sqrt{\sigma_{\mathbf{v}}^2 \frac{r_1^* r_2^*}{\rho_1^* \rho_2^*}} \frac{(1 - \rho_1 z^{-1})(1 - \rho_2 z^{-1})}{(1 - r_1 z^{-1})(1 - r_2 z^{-1})}.$$

And the causal filter $h = (h_n, n > -\infty)$ is

$$\begin{aligned} h_n &= 0 && \text{if } n < 0 \\ h_n &= \phi(0) && \text{if } n = 0 \\ h_n &= \phi(1) - \phi(0)(r_1 + r_2) && \text{if } n = 1 \\ h_n &= \phi(n) - (r_1 + r_2)\phi(n-1) + r_1 r_2 \phi(n-2) && \text{if } n \geq 2 \end{aligned}$$

Where

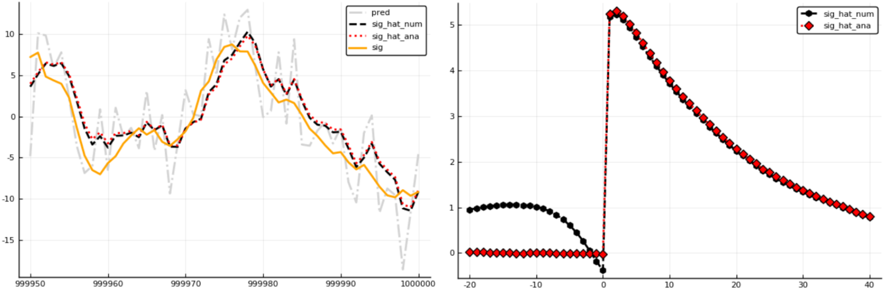
$$\phi(n) = \phi_0^2 \sum_{k=0}^{n} \gamma(n-k)\alpha(k), \qquad \alpha(n) = \sum_{k=0}^{n} \rho_1^{n-k}\rho_2^k, \qquad \beta(n) = \sum_{k=0}^{n} r_1^{n-k} r_2^k,$$

and

$$\gamma(n) = \begin{cases} \sum_{k=0}^{\infty} \alpha^*(k-n)\beta(k) & n \leq 0 \\ \sum_{k=0}^{\infty} \alpha^*(k)\beta(k+n) & n > 0 \end{cases}$$

Example 3: AR(2)) Signal, Additive WN

Here is a run, with $r1, r2 = .5, .95$, $\sigma_{\mathbf{v}} = 4$. The trajectory has 10^6 steps after discarding 10^3 steps.



Left: A window of the time series for the signal (orange), the predictors (light gray), the estimated signal using the analytic and numerical Wiener filter (red, black).

Right: The covariance between errors (red from analytic filter, black from numerical) and predictors (observations).

Example 4: AR(2) Signal, Filtered, Ad. WN

Let us consider again the stationary autoregressive process of order 2,

$$y_n = (r_1 + r_2)y_{n-1} - r_1 r_2 y_{n-2} + u_n, \qquad \text{for } n > -\infty$$

for $r_1, r_2 \in \{z : |z| < 1\}$. This time however, we define the observations to be the signal y operated upon by a finite impulse response time invariant filter w with additive white noise.

$$x_n = (w * y)_n + v_n, \qquad \text{for } n > -\infty.$$

For simplicity let

$$w = (\dots, 0, \boxed{1}, w_1, w_2, 0, \dots),$$

where the box indicate the element indexed by 0 and $w_1, w_2 \in \mathbb{R}$, then write

$$W(z) = \sum_{k=-\infty}^{\infty} w_k z^{-k} = 1 + w_1 z^{-1} + w_2 z^{-2}.$$

Observe that

$$S_{\mathbf{yx}}(z) = S_{\mathbf{y}}(z)W^*(z^{-*}) \qquad \text{and} \qquad S_{\mathbf{x}} = W(z)S_{\mathbf{y}}(z)W^*(z^{-*}) + \sigma_{\mathbf{v}}^2.$$

So,

$$S_{\mathbf{yx}}(z) = \frac{1 + w_1 z + w_2 z^2}{(1 - r_1 z^{-1})(1 - r_2 z^{-1})(1 - r_1^* z)(1 - r_2^* z)}$$

and

$$S_{\mathbf{x}}(z) = \frac{w_2 + \sigma_{\mathbf{v}}^2 r_1^* r_2^*}{\rho_1^* \rho_2^*} \cdot \frac{(1 - \rho_1 z^{-1})(1 - \rho_2 z^{-1})(1 - \rho_1^* z)(1 - \rho_2^* z)}{(1 - r_1 z^{-1})(1 - r_2 z^{-1})(1 - r_1^* z)(1 - r_2^* z)}$$

For suitable ρ_1, ρ_2 (which depend on w_1, w_2)

So, the causal filter $h = (h_n, n > -\infty)$ is

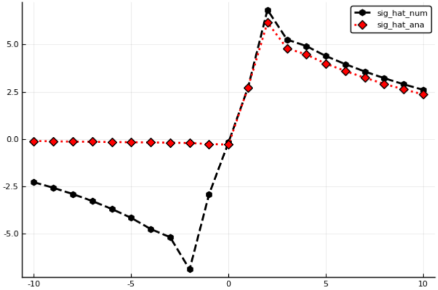
$$\begin{aligned} h_n &= 0 && \text{if } n < 0 \\ h_n &= \psi(0) && \text{if } n = 0 \\ h_n &= \psi(1) - \psi(0)(r_1 + r_2) && \text{if } n = 1 \\ h_n &= \psi(n) - (r_1 + r_2)\psi(n-1) + r_1 r_2 \psi(n-2) && \text{if } n \geq 2 \end{aligned}$$

Where

$$\psi(n) = \psi_0^2 \sum_{k=0}^{n} \left[\gamma(n-k) + w_1\gamma(n-k+1) + w_2\gamma(n-k+2)\right]\alpha(k).$$

Example 4: AR(2) Signal, Filtered, Ad. WN

Here is a run, with $r_1, r_2 = -0.2, 0.9$, $w_1, w_2 = -0.1, 5$, and $\sigma_{\mathbf{v}} = 1.1$. The trajectory has 10^6 steps after discarding 10^3 steps.



Left: A window of the time series for the signal (orange), the predictors (light gray), the estimated signal using the analytic and numerical Wiener filter (red, black).

Right: The covariance between errors (red from analytic filter, black from numerical) and predictors (observations).

Thank you!