$f(e)$	$Ef(e)$
e^Te	nM_2
ee^Te	$M_3 1_n$
e^Tee^Te	$nM_4 + n(n-1)M_2^2$

Table 1: The moments of e. table:moment-e

1 Condition and Property of MISQ Filter

ition-property-MISQ-filter

1.1 Modeling Time Series

sec:modeling-time-series

We follow the model of time series in the literature. Suppose $X = (X_1, X_2, \cdots, X_n)^T$ is the underlying random vector of observed time series, we assume

$$X = \mu + e, \tag{1}$$

eq:model-X

where $\mu = (\mu_1, \mu_2, \cdots, \mu_n)^T \in \Re^n$ and $e = (e_1, e_2, \cdots, e_n)^T$ is a zero-mean random vector.

For e, we assume that e has following properties:

- $e_1, e_2, ..., e_n$ are identical and independent distributed (i.i.d.). We denote the distribution of e_i as ε. This assumption makes our work much more easily and is common in the literature such as
- $E\varepsilon = 0$.
- $E\varepsilon^k$ are existed and bounded for $k = 2, 3, 4$.

The moments of e are listed in Table moment-e ??.

For μ, we assume that μ is continuous which is expressed mathematically as:

$$\mu_i \approx \mu_{i+1}. \tag{2}$$

eq:model-mu

Many readers might not agree with this assumption, but it is existed in many existed work and needed to handle time series without multiple observations for each time stamp.

Moving average is a popular techniques to estimate the μ in model eq:model-X 1:

$$\begin{aligned}
\hat{\mu_i} &= \frac{1}{2h+1} \sum_{k=-h}^{h} X_{i+k} \\
&= \frac{1}{2h+1} \sum_{k=-h}^{h} \mu_{i+k} + e_{i+k} \\
&= \frac{1}{2h+1} \sum_{k=-h}^{h} \mu_{i+k} + \frac{1}{2h+1} \sum_{k=-h}^{h} e_{i+k}. \qquad (3)
\end{aligned}$$

Under the i.i.d. assumption of e_i, the variance is reduced from M_2 to $\frac{1}{2h+1} M_2$, so $\hat{\mu_i}$ might be more closer to μ compared to X_i if $\mu_i \approx \frac{1}{2h+1} \sum_{k=-h}^{h} \mu_{i+k}$. Therefore, the reader should notice that moving average also assumes Eq eq:model-mu 2 intrinsically.

1.2 Linear Filtered Time Series

In this paper, we define the linear filtered time series $\tilde{X}$ as

$$\tilde{X}_i = \begin{cases} \sum_{k=0}^{m} \phi_k X_{i+k} & \text{if } i \geq 1 \text{ and } i \leq n-m \\ 0 & otherwise \end{cases} \tag{4}$$

eq:filter-X

We extend the range of subscribe for convenience.

If $\phi = (-1, 1)^T$, the filtered time series is the one in MISQ ??. If $\phi = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5})$, the filtered time series is the moving averaged time series.

To write the formula in matrix form, we let Φ be a $(n-m) \times n$ matrix with i-th row and j-th column entries

$$\Phi_{i,j} = \phi_{j-i}. \tag{5}$$

eq:definition-Phi_ij

Therefore, the filtered time series is ΦX.

Here, we add a constraint to ϕ:

$$\sum_{k=0}^{m} \phi_k = 0. \tag{6}$$

eq:model-phi

When $m << n$, we combine the Eq eq:model-mu 2 and Eq eq:model-phi 6:

$$\begin{aligned} \Phi X &= \Phi\mu + \Phi e \quad \text{according to Eq eq:model-X 1} \\ &\approx \Phi e. \end{aligned} \tag{7}$$

eq:model-Phi-X

The last approximation holds because of

$$\begin{aligned} \sum_{k=0}^{m} \phi_{i+k}\mu_{i+k} &\approx \sum_{k=0}^{m} \phi_{i+k}\mu_i \\ &= 0. \end{aligned} \tag{8}$$

Therefore, we have an approximation of Φe. To make inference according to Φe, we list the moments of Φe in Table table:moment-Phie 2. We derive moment estimators for M_2, M_3, and M_4.

Note that we define a $n \times n$ matrix $\Lambda = \Phi^T \Phi$ with its i-th row and j-th column entry $\Lambda_{i,j}$ for convenience. The vector $diag(\Lambda) = (\Lambda_{1,1}, \Lambda_{2,2}, \cdots, \Lambda_{n,n})^T \in \Re^n$ is the diagonal vector of Λ. The $tr(\Lambda)$ is the trace of Λ. Here, we give two properties of Λ.

property:sum-diag-Lambda **Property 1**

$$\sum_{i=1}^{n} \Lambda_{i,i} = (n-m) \sum_{k=0}^{m} \phi_k^2. \tag{9}$$

Proof

$f(e)$	$Ef(e)$
$e^T\Phi^T\Phi e$	$M_2(n-m)\sum_{k=0}^{m}\phi_k^2$
$ee^T\Phi^T\Phi e$	$M_3 diag(\Lambda)$
$e^Tee^T\Phi^T\Phi e$	$(M_4+(n-1)M_2^2)(n-m)\sum_{k=0}^{m}\phi_k^2$
$e^T\Phi^T\Phi ee^T\Phi^T\Phi e$	$(M_4-3M_2^2)(\sum_i \Lambda_{i,i}^2)+M_2^2(n-m)^2(\sum_{i=0}^{n}\Lambda_{i,i})^2+2M_2^2(\sum_{i,j}\Lambda_{i,j}^2)$

Table 2: The moments of e. table:moment-Phie

We directly expand the definition of $\Lambda_{i,i}$ as follow:

$$\begin{aligned}\sum_{i=1}^{n}\Lambda_{i,i} &= \sum_{i=1}^{n}\sum_{k=1}^{n-m}\Phi_{k,i}^2\\ &= \sum_{i=1}^{n}\sum_{k=1}^{n-m}\phi_{i-k}^2\\ &= \sum_{k=1}^{n-m}\sum_{i=k}^{k+m}\phi_{i-k}^2\\ &= (n-m)\sum_{k=0}^{m}\phi_k^2\end{aligned}$$

Property property:sum-diag-Lambda 1 will be used to evaluate the variance at the Subsection sec:misq-estimator-variance 1.3.

property:diag-Lambda-equa **Property 2** *If $i > m$ and $i \le n-m$,*

$$\Lambda_{i,i} = \sum_{k=0}^{m}\phi_k^2. \tag{10}$$

Proof

Expand the definition of $\Lambda_{i,i}$ to prove this property.

The reader should notice that Eq eq:model-Phi-X 7 is more likely to be true in practice if m is small.

1.3 MISQ distance estimator and variance

sec:misq-estimator-variance

If there are two time series S_1 and S_2 in model eq:model-X 1 and model eq:model-mu 2, let $X = S_1 - S_2$ is still in model eq:model-X 1 and model eq:model-mu 2. The sum of square of $\mu^T\mu$ is the distance between S_1 and S_2 without errors. For better inference, we discuss the estimation of $\frac{\mu^T\mu}{n}$ called *mean distance* in this paper.

We define the estimator of mean distance corresponding to ϕ as

$$\hat{D}_\phi(X) = \frac{X^TX}{n} - \frac{X^T\Phi^T\Phi X}{(n-m)\sum_{k=0}^{m}\phi_m^2}, \tag{11}$$

because its expectation is

$$
\begin{aligned}
E\hat{D}_\phi(X) &\approx \frac{\mu^T\mu + Ee^Te}{n} - \frac{Ee^T\Phi^T\Phi e}{(n-m)\sum\limits_{k=0}^{m}\phi_m^2} \\
&= \frac{\mu^T\mu + nM_2}{n} - \frac{(n-m)M_2\sum\limits_{k=0}^{m}\phi_m^2}{(n-m)\sum\limits_{k=0}^{m}\phi_m^2} \\
&= \frac{\mu^T\mu}{n} \qquad (12) \quad \boxed{\texttt{eq:hatD}}
\end{aligned}
$$

The variance of $\hat{D}_\phi(X)$ is

$$
\begin{aligned}
& Var(\hat{D}_\phi(X)) \\
= & \frac{1}{n^2}Var(X^TX) - 2\frac{1}{n(n-m)\sum\limits_{k=0}^{m}\phi_k^2}Cov(X^TX, X^T\Phi^T\Phi X) \\
+ & \frac{1}{(n-m)^2\left(\sum\limits_{k=0}^{m}\phi_k^2\right)^2}Var(X^T\Phi^T\Phi X) \\
= & \frac{1}{n^2}\left(4M_2\mu^T\mu + 4M_3\mu^T1_n + nM_4 - nM_2^2\right) \\
- & \frac{2}{n(n-m)\sum\limits_{k=0}^{m}\phi_k^2}\left(2M_3\mu^T diag(\Lambda) + (M_4 - M_2^2)(n-m)\sum_{k=0}^{m}\phi_k^2\right) \\
+ & \frac{1}{(n-m)^2\left(\sum\limits_{k=0}^{m}\phi_k^2\right)^2}\left((M_4 - 2M_2^2)\sum_{i=1}^{n}\Lambda_{i,i}^2 + M_2^2\sum_{i,j}\Lambda_{i,j}^2\right) \\
= & \left(\frac{4\mu^T\mu}{n^2}\right)M_2 + \left(\frac{4\mu^T1_n}{n^2} - \frac{4\mu^T diag(\Lambda)}{n(n-m)\sum\limits_{k=0}^{m}\phi_k^2}\right)M_3 \\
+ & \left(\frac{\sum\limits_{k=1}^{n}\Lambda_{i,i}^2}{(n-m)^2(\sum\limits_{k=0}^{m}\phi_k^2)^2} - \frac{1}{n}\right)M_4 + \left(\frac{1}{n} - \frac{\sum\limits_{i,j}\Lambda_{i,j} - 2\sum\limits_{i=1}^{n}\Lambda_{i,i}^2}{(n-m)^2(\sum\limits_{k=0}^{m}\phi_k^2)^2}\right)M_2^2 (13)
\end{aligned}
$$

1.4 Estimable

According to Table table:moment 2 and Eq eq:hatD 12, we can replace some unobserved quantity such as $M_2, M_3, M_4, \mu^T\mu$ with their moment estimator. However, μ^T1_n and

$\mu^T diag(\Lambda)$ have no related moment estimator, so we must make sure that the coefficient corresponding to them are negliable, i.e.

$$ \tag{14} $$