

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

Table 1: The moments of  $e$ .

table:moment-e

## 1 Condition and Property of MISQ Filter

### 1.1 Modeling Time Series

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

$$X = \mu + e, \quad (1)$$

eq:model-X

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

For  $e$ , we assume that  $e$  has following properties:

- $e_1, e_2, \dots, e_n$  are identical and independent distributed (i.i.d.). We denote the distribution of  $e_i$  as  $\varepsilon$ . 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  $\mu$ , we assume that  $\mu$  is continuous which is expressed mathematically as:

$$\mu_i \approx \mu_{i+1}. \quad (2)$$

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  $\mu$  in model [eq:model-X](#):

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

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  $\mu$  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](#) intrinsically.

## 1.2 Linear Filtered Time Series

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

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

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  $\Phi$  be a  $(n-m) \times n$  matrix with  $i$ -th row and  $j$ -th column entries

$$\Phi_{i,j} = \phi_{j-i}. \quad (5) \quad \text{eq:definition-Phi-ij}$$

Therefore, the filtered time series is  $\Phi X$ .

Here, we add a constraint to  $\phi$ :

$$\sum_{k=0}^m \phi_k = 0. \quad (6) \quad \text{eq:model-phi}$$

When  $m \ll n$ , we combine the Eq [eq:model-phi](#) and Eq [eq:model-phi](#):

$$\begin{aligned} \Phi X &= \Phi \mu + \Phi e \quad \text{according to Eq [eq:model-X](#)} \\ &\approx \Phi e. \end{aligned} \quad (7) \quad \text{eq:model-Phi-X}$$

The last approximation holds because of

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

Therefore, we have an approximation of  $\Phi e$ . To make inference according to  $\Phi e$ , we list the moments of  $\Phi e$  in [Table 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  $\text{diag}(\Lambda) = (\Lambda_{1,1}, \Lambda_{2,2}, \dots, \Lambda_{n,n})^T \in \mathbb{R}^n$  is the diagonal vector of  $\Lambda$ . The  $\text{tr}(\Lambda)$  is the trace of  $\Lambda$ . Here, we give two properties of  $\Lambda$ .

property:sum-diag-Lambda

### Property 1

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

### Proof

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

Table 2: The moments of  $e$ .

table:moment-Phi

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

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

Property [property:sum-diag-Lambda](#) 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  $i \leq n - m$ ,

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

**Proof**

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

The reader should notice that Eq [eq:model-Phi-X](#) is more likely to be true in practice if  $m$  is small.

ec:misq-estimator-variance

### 1.3 MISQ distance estimator and variance

If there are two time series  $S_1$  and  $S_2$  in model [eq:model-X](#) and model [eq:model-mu](#) 2, let  $X \equiv S_1 - S_2$  is still in model [eq:model-X](#) 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  $\phi$  as

$$\hat{D}_\phi(X) = \frac{X^T X}{n} - \frac{X^T \Phi^T \Phi X}{(n-m) \sum_{k=1}^m \phi_k^2}, \quad (11)$$

because its expectation is

$$\begin{aligned} E\hat{D}_\phi(X) &\approx \frac{\mu^T \mu + Ee^T e}{n} - \frac{Ee^T \Phi^T \Phi e}{(n-m) \sum_{k=1}^m \phi_k^2} \\ &= \frac{\mu^T \mu + nM_2}{n} - \frac{(n-m)M_2 \sum_{k=1}^m \phi_k^2}{(n-m) \sum_{k=1}^m \phi_k^2} \\ &= \frac{\mu^T \mu}{n} \end{aligned} \quad (12) \quad \boxed{\text{eq:hatD}}$$

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

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

## 1.4 Estimable

According to Table [table:moment-Phi-hatD](#) and Eq 12, we can replace some unobserved quantity such as  $M_2, M_3, M_4, \mu^T \mu$  with their moment estimator. However,  $\mu^T 1_n$  and  $\mu^T \text{diag}(\Lambda)$  have no related moment estimator, so we must make sure that the coefficient corresponding to them are negligible, i.e.

(14)