Variance of Stepwise Sample Means

Jingxuan Yang*

August 19, 2022

In this note, we will analyze the variance of stepwise sample means.

Suppose the data samples are $\{x_i\}_{i=1}^n$. The sample mean is

$$\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{1}$$

The sample variance is

$$\sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}_n^2. \tag{2}$$

Therefore, the matrix form of sample variance is

$$\begin{aligned}
\sigma_n^2 &= \frac{1}{n}(x_1^2 + \cdots + x_n^2) - \frac{1}{n^2}(x_1 + \cdots + x_n)^2 \\
&= \frac{1}{n^2} x^\top \begin{bmatrix} n & & & \\ & n & & \\ & & \ddots & \\ & & & n \end{bmatrix} x - \frac{1}{n^2} x^\top \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} x \\
&= \frac{1}{n^2} x^\top \begin{bmatrix} n-1 & -1 & \cdots & -1 \\ -1 & n-1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & n-1 \end{bmatrix} x \\
&\triangleq \frac{1}{n^2} x^\top V x,
\end{aligned} \tag{3}$$

where V is defined as the variance matrix of the sample mean. It's a zero-sum matrix, i.e., the sum of all elements equals zero.

Denote the mean of stepwise sample means $\{\bar{x}_j\}_{j=1}^n$ as

$$\bar{\bar{x}}_n = \frac{1}{n}\sum_{j=1}^{n} \bar{x}_j = \frac{1}{n}\sum_{j=1}^{n}\frac{1}{j}\sum_{i=1}^{j} x_i = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=i}^{n}\frac{1}{j} x_i = \frac{1}{n}\sum_{i=1}^{n} \gamma_{i,n} x_i, \tag{4}$$

*Jingxuan Yang is with the Department of Automation, Tsinghua University, Beijing 100084, China (email: yangjx20@mails.tsinghua.edu.cn).

where

$$\gamma_{i,n} \triangleq \sum_{j=i}^{n} \frac{1}{j}, \quad \frac{1}{n}\sum_{i=1}^{n} \gamma_{i,n} = 1, \tag{5}$$

and the variance of stepwise sample means as

$$\bar{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n} (\bar{x}_i - \bar{\bar{x}}_n)^2 = \frac{1}{n}\sum_{i=1}^{n} \bar{x}_i^2 - \bar{\bar{x}}_n^2, \tag{6}$$

which is actually the variance of the mean of stepwise sample means. There are two mean operations here, and thus this variance is usually less than the sample variance.

The matrix form of this variance is

$$\begin{aligned}
\bar{\sigma}_n^2 &= \frac{1}{n}\left[x_1^2 + \left(\frac{x_1+x_2}{2}\right)^2 + \cdots + \left(\frac{x_1+\cdots+x_n}{n}\right)^2\right] - \frac{1}{n^2}\left(\sum_{i=1}^{n}\gamma_{i,n}x_i\right)^2 \\
&= \frac{1}{n^2} x^\top \begin{bmatrix} n\xi_{1,n} & n\xi_{2,n} & \cdots & n\xi_{n,n} \\ n\xi_{2,n} & n\xi_{2,n} & \cdots & n\xi_{n,n} \\ \vdots & \vdots & \ddots & \vdots \\ n\xi_{n,n} & n\xi_{n,n} & \cdots & n\xi_{n,n} \end{bmatrix} x \\
&\quad - \frac{1}{n^2} x^\top \begin{bmatrix} \gamma_{1,n}^2 & \gamma_{1,n}\gamma_{2,n} & \cdots & \gamma_{1,n}\gamma_{n,n} \\ \gamma_{2,n}\gamma_{1,n} & \gamma_{2,n}^2 & \cdots & \gamma_{2,n}\gamma_{n,n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n,n}\gamma_{1,n} & \gamma_{n,n}\gamma_{2,n} & \cdots & \gamma_{n,n}^2 \end{bmatrix} x \\
&= \frac{1}{n^2} x^\top \begin{bmatrix} n\xi_{1,n} - \gamma_{1,n}^2 & n\xi_{2,n} - \gamma_{1,n}\gamma_{2,n} & \cdots & n\xi_{n,n} - \gamma_{1,n}\gamma_{n,n} \\ n\xi_{2,n} - \gamma_{2,n}\gamma_{1,n} & n\xi_{2,n} - \gamma_{2,n}^2 & \cdots & n\xi_{n,n} - \gamma_{2,n}\gamma_{n,n} \\ \vdots & \vdots & \ddots & \vdots \\ n\xi_{n,n} - \gamma_{n,n}\gamma_{1,n} & n\xi_{n,n} - \gamma_{n,n}\gamma_{2,n} & \cdots & n\xi_{n,n} - \gamma_{n,n}^2 \end{bmatrix} x \\
&\triangleq \frac{1}{n^2} x^\top \bar{V} x,
\end{aligned} \tag{7}$$

where $\xi_{i,n} \triangleq \sum_{j=i}^{n} \frac{1}{j^2}$, and $\bar{V}$ is the variance matrix of the stepwise sample means.

Taking the first element of $\bar{V}$ as an example, it can be expressed as

$$\begin{aligned}
n\xi_{1,n} - \gamma_{1,n}^2 &= n\sum_{j=1}^{n}\frac{1}{j^2} - \left(\sum_{j=1}^{n}\frac{1}{j}\right)^2 \\
&= n\left(1 + \frac{1}{2^2} + \cdots \frac{1}{n^2}\right) - \left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right)^2 \\
&\to n\frac{\pi^2}{6} - (\ln n + \gamma)^2, \quad n \to \infty,
\end{aligned} \tag{8}$$

where $\gamma \approx 0.5772$ is the Euler number. It can be found that this element will exceed $n-1$ more and more as n increases. For other elements, we study them by experiments.

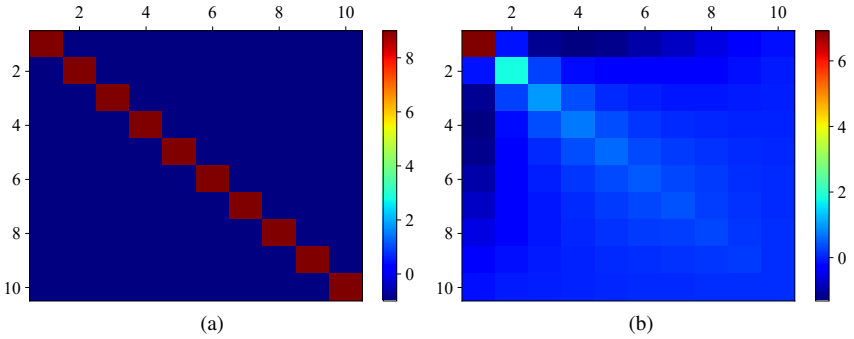

Figure 1: (a) Variance matrix of samples for $n = 10$; (b) Variance matrix of stepwise sample means for $n = 10$.

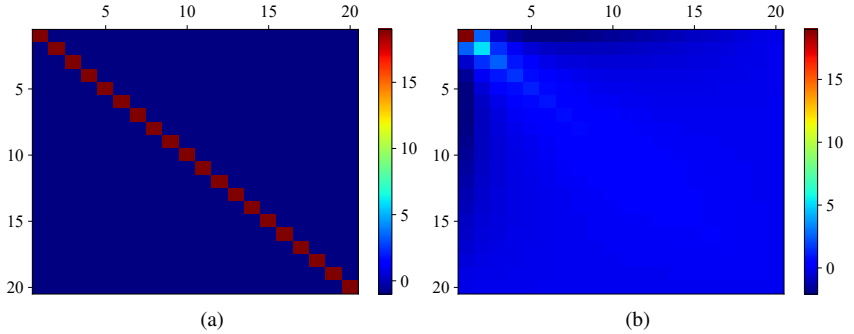

Figure 2: (a) Variance matrix of samples for $n = 20$; (b) Variance matrix of stepwise sample means for $n = 20$.

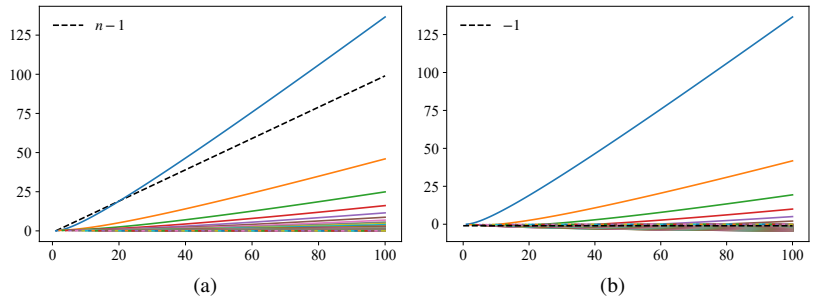

Figure 3: (a) The diagonal elements of $\bar{V}$; (b) The first column of $\bar{V}$.

As illustrations, the variance matrices of samples and stepwise sample means for $n = 10$ and $n = 20$ are shown in Fig. 1 and Fig. 2, respectively. Furthermore, the

diagonal elements and the first column of $\bar{V}$ with $n = 1, 2, \ldots, 100$ are presented in Fig. 3a and Fig. 3b, respectively. It can be seen that $\bar{V}$ puts more weights to the front samples and less to the rear samples, while V lays equal weights on each sample. Therefore, the arrangements of the samples make less difference on the variance of stepwise sample means than on the variance of samples.

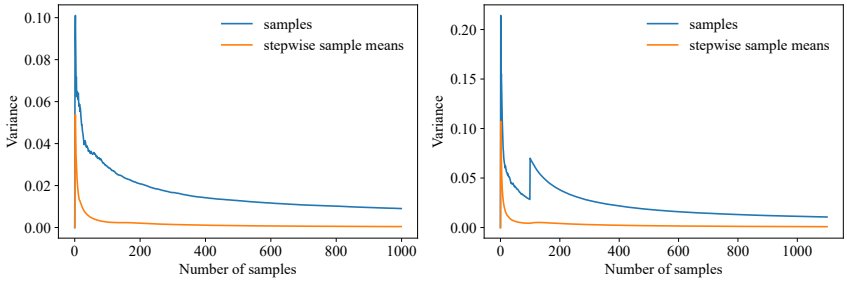

Figure 4: (a) Variance of samples and stepwise sample means without outliers; (b) Variance of samples and stepwise sample means with an outlier.

For sample variance, the poor performance of some front samples will require many samples to mitigate. Specifically, a poor sample will bring huge variance at front, while can be ignored if in the rear. This phenomenon is demonstrated in Fig. 4, where an outlier is inserted as the 101^{st} sample. The outlier brings a sharp variance increase to the sample variance, while throwing no impact on the variance of stepwise sample means.