Liet. matem. rink., **47**, *spec. nr.*, 2007, 411–416

On recursive expressions for statistics of decimated sequences

Rimantas PUPEIKIS (MII)
e-mail: pupeikis@ktl.mii.lt

Abstract. In what follows we introduce the recursive approach for calculating statistical moments of decimated realizations. We prove the corollaries referring to recursive calculation and present an example for any realization of 17 sample.

Keywords: discrete-time signal, sampling, decimation, statistics.

1. Introduction

While processing discrete-time signals (sequences) there arises a problem to retrieve maximal information as well as to reduce the amount of calculations on samples. In such a case, the data decimation by means of a downsampling operation is used [8]. This is a time-scaling operation that is equivalent to changing the sampling rate F_s of an analogous signal from $1/T_s$ to $1/2T_s$, where T_s is a sampling period and its reciprocal $1/T_s = F_s$, i.e., is to decreasing the sampling rate two times. In such a case, the number of samples to be processed decreases twice, as well. In general, the basic sampling frequency F_s could be decreased by the fixed integer number of times. It is known that the decimation process ought to be stopped before the frequency content of the signal is above the new Nyquist frequency F_N. The sampling rate determination techniques are proposed for dicrete process control and identification of dynamic systems [5, 1, 7, 4, 11], as well as for improving spectral resolution while solving a spectral estimation problem [3, 6, 9, 10]. On the other hand, it is important, first of all, to calculate simple statistical characteristics of decimated realizations and a basic non-decimated discrete-time signal. There also arises a problem to reduce the number of operations needed for calculations of first and second statistical moments, without storing decimated realizations in a memory of a computer. Here both problems have been solved, using the recursive approach proposed in the paper.

2. Statement of the problem

Assume that we consider a discrete-time signal $U(kT_s)\ \forall\, k \in \overline{0, \mathrm{N}}$ that is obtained by sampling its continuous-time counterpart $U(t)$ with sampling frequency F_s. Here N is the general number of samples of the basic signal $U(kT_s)\ \forall\, k \in \overline{0, \mathrm{N}}$ under consideration; t is a continuous time variable. Suppose, for simplicity, that N is divisible n times by 2, i.e., $\mathrm{N} = 2^n$.

After a multiplex decimation of the realization $u(kT_s)\ \forall k \in \overline{0,\mathrm{N}}$, one has the set Ω of such sequences $x_1(k) \equiv u(kT_s)\ \forall\ k \in \overline{0,\mathrm{N}},\ x_2(k) \equiv u(k2T_s)\ \forall k \in \overline{0,\mathrm{N}/2},\ x_3(k) \equiv u(k4T_s)\ \forall\ k \in \overline{0,\mathrm{N}/4}, \ldots,\ x_{n-1}(k) \equiv u(knT_s/2)\ \forall k \in \overline{0,2\mathrm{N}/n},\ x_n(k) \equiv u(knT_s)\ \forall\ k \in \overline{0,\mathrm{N}/n}$ of the same signal $U(kT_s)\ \forall\ k \in \overline{0,\mathrm{N}}$. Afterwards, the first and second statistical moments could be calculated by processing different sequences from the set Ω, respectively, using well-known formulas [2]. Two different sequences are finally obtained containing mean and variance values for each of the above mentioned realizations from the set Ω that have been stored in the memory of a computer.

The aim of the given paper is: firstly, to calculate the abovementioned statistics without storing decimated sequences in the memory of a computer, secondly, to reduce the number of calculations by applying the recursive approach that will use the information obtained by processing the previous decimated realization from the same set Ω in the current operation.

3. Recursive expressions for mean and variance values

In order to work out the recursive expressions for the first and second central order statistics, let us formulate statements on the calculation of means and variances of the decimated realizations as well as non-decimated ones.

Corollary 1. The mean of each realization $m(\cdot)$ from the set Ω is calculated using the recursive expression of the form

$$m(iT_s) = \frac{i}{\mathrm{N}+i}\left\{\frac{N+2i}{2i}m(2iT_s) + \sum_{l=1}^{\mathrm{N}/(2i)} u\big(iT_s(2l-1)\big)\right\} \tag{1}$$

$\forall\ i = 1, 2, 4, 8, \ldots, n/2$. Here $m(iT_s)$ and $m(i2T_s)\ \forall\ i = 1, 2, 4, 8, \ldots, n/2$ are mean values of the current and previous iterations.

Corollary 2. The variance of each realization $var(\cdot)$ from the set Ω is calculated using the recursive expression of the form

$$var(iT_s) = \frac{i}{\mathrm{N}+i}\left\{\frac{N+2i}{2i}var(2iT_s) + \sum_{l=1}^{\mathrm{N}/(2i)} \dot{u}^2\big(iT_s(2l-1)\big)\right\} \tag{2}$$

$\forall\ i = 1, 2, 4, 8, \ldots, n/2$. Here

$$\dot{u}\big(iT_s(2l-1)\big) = u\big(iT_s(2l-1)\big) - m\big(iT_s(2l-1)\big) \tag{3}$$

$\forall\ i = 1, 2, 4, 8, \ldots, n/2$. Here $var(iT_s)$ and $var(i2T_s)\ \forall\ i = 1, 2, 4, 8, \ldots, n/2$ are variance values of the current and previous iterations.

Proof of Corollary 1. Let us describe now the signal $u(kT_s)\ \forall\ k \in \overline{0,\mathrm{N}}$ resolving it into a sum of unit sample sequences $\delta(kT_s - lT_s)$ such as [8]

$$u(kT_s) = \sum_{l=0}^{\mathrm{N}} u(lT_s)\delta\big[T_s(k-l)\big] \tag{4}$$

because $\delta(kT_s - lT_s)$ is zero everywhere except at $k = l$, where its value is a unity. The same could be rewritten as the sum of two sequences

$$u(kT_s) = \sum_{l=0}^{\mathrm{N}/2} u(l2T_s)\delta[T_s(k-2l)] + \sum_{l=0}^{\mathrm{N}/2} u(l2T_s)\delta\big[T_s(k-2l-1)\big]. \qquad (5)$$

On the righthandside of expression (5), the first sequence is

$$\sum_{l=0}^{\mathrm{N}/2} u(l2T_s)\delta\big[T_s(k-2l)\big]$$
$$= \sum_{l=0}^{\mathrm{N}/4} u(l4T_s)\delta\big[T_s(k-4l)\big] + \sum_{l=0}^{\mathrm{N}/4} u(l4T_s)\delta\big[T_s(k-4l-2)\big]. \qquad (6)$$

On the other hand,

$$\sum_{l=0}^{\mathrm{N}/4} u(l4T_s)\delta\big[T_s(k-4l)\big]$$
$$= \sum_{l=0}^{\mathrm{N}/8} u(l8T_s)\delta\big[T_s(k-8l)\big] + \sum_{l=0}^{\mathrm{N}/8} u(l8T_s)\delta\big[T_s(k-8l-4)\big]. \qquad (7)$$

By proceeding with this, we can finally obtain

$$\sum_{l=0}^{2\mathrm{N}/n} u(lnT_s/2)\delta\big[T_s(k-nl/2)\big]$$
$$= \sum_{l=0}^{\mathrm{N}/n} u(lnT_s)\delta\big[T_s(k-nl)\big] + \sum_{l=0}^{\mathrm{N}/n} u(lnT_s)\delta\big[T_s(k-nl-n/2)\big]. \qquad (8)$$

It follows from (8) that the mean of the last decimated sequence $x_n(k)\ \forall\, k \in \overline{0, \mathrm{N}/n}$ is

$$m(nT_s) = \frac{n}{\mathrm{N}+n} \sum_{l=0}^{\mathrm{N}/n} u(lnT_s), \qquad (9)$$

while that of the previous decimated sequence $x_{n-1}(k)\ \forall\, k \in \overline{0, 2\mathrm{N}/n}$ is

$$m(nT_s/2) = \frac{n}{2\mathrm{N}+n}\left\{\frac{\mathrm{N}+n}{n} m(nT_s) + \sum_{l=1}^{\mathrm{N}/n} u\big(T_s n(l-1/2)\big)\right\}. \qquad (10)$$

Continuing with this in the reverse order one could obtain the recursive formulas for calculating means. In view of (5) and (6) for the first-decimated realization and for the basic one, those recursive formulas are

$$m(2T_s) = \frac{2}{\mathrm{N}+2}\left\{\frac{\mathrm{N}+4}{4}m(4T_s) + \sum_{l=1}^{\mathrm{N}/4} u\big(2T_s(2l-1)\big)\right\} \tag{11}$$

and

$$m(T_s) = \frac{1}{\mathrm{N}+1}\left\{\frac{\mathrm{N}+2}{2}m(2T_s) + \sum_{l=1}^{\mathrm{N}/2} u\big(T_s(2l-1)\big)\right\}, \tag{12}$$

respectively. Thus, the general expression for calculating means is of the form (1).

The proof of the Corollary 2 is similar as that of the Corollary 1.

4. Example

Let us now calculate mean and variance values having any realization of the basic non-decimated discrete-time signal $U(kT_s)\ \forall k \in \overline{0,\mathrm{N}}$ with N=16 and arbitrary T_s. After decimating this realization, we get the set Ω of realizations: $u(kT_s)\ \forall\ k \in \overline{0,16}$, $u(k2T_s)\ \forall k \in \overline{0,8}$, $u(k4T_s)\ \forall\ k \in \overline{0,4}$, $u(k8T_s)\ \forall k \in \overline{0,2}$. First of all, by substituting n=8 in formulas (1), (2), we calculate the mean and variance values of the last decimated realization:

$$m(8T_s) = \frac{8}{16+8}\sum_{l=0}^{2} u(8lT_s) = \frac{1}{3}\big[u(0) + u(8T_s) + u(16T_s)\big], \tag{13}$$

$$var(8T_s) = \frac{8}{16+8}\sum_{l=0}^{2} \dot{u}^2(8lT_s) = \frac{1}{3}\big[\dot{u}^2(0) + \dot{u}^2(8T_s) + \dot{u}^2(16T_s)\big], \tag{14}$$

respectively. Then the mean and variance values of the second decimated realization $u(k4T_s)\ \forall\ k \in \overline{0,4}$ are calculated according to recursive expressions (1),(2) as follows:

$$\begin{aligned} m(4T_s) &= \frac{4}{16+4}\left\{\frac{16+8}{8}m(8T_s) + \sum_{l=1}^{2} u\big(4T_s(2l-1)\big)\right\} \\ &= \frac{1}{5}\big[3m(8T_s) + u(4T_s) + u(12T_s)\big], \end{aligned} \tag{15}$$

$$\begin{aligned} var(4T_s) &= \frac{4}{16+4}\left\{\frac{16+8}{8}var(8T_s) + \sum_{l=1}^{2} \dot{u}^2\big(4T_s(2l-1)\big)\right\} \\ &= \frac{1}{5}\big[3var(8T_s) + \dot{u}^2(4T_s) + \dot{u}^2(12T_s)\big]. \end{aligned} \tag{16}$$

Here $\dot{u}(4T_s) = u(4T_s) - m(4T_s)$ and $\dot{u}(12T_s) = u(12T_s) - m(4T_s)$. Using the same expressions (1),(2), we get the mean of the first decimated realization $u(k2T_s)$

$\forall\, k \in \overline{0, 8}$

$$m(2T_s) = \frac{2}{16+2}\left\{\frac{16+4}{4}m(4T_s) + \sum_{l=1}^{4} u\big(2T_s(2l-1)\big)\right\}$$
$$= \frac{1}{9}\big[5m(4T_s) + u(2T_s) + u(6T_s) + u(10T_s) + u(14T_s)\big] \qquad (17)$$

and variance

$$var(2T_s) = \frac{2}{16+2}\left\{\frac{16+4}{4}var(4T_s) + \sum_{l=1}^{4} \dot{u}^2\big(2T_s(2l-1)\big)\right\}$$
$$= \frac{1}{9}\big[5var(4T_s) + \dot{u}^2(2T_s) + \dot{u}^2(6T_s) + \dot{u}^2(10T_s) + \dot{u}^2(14T_s)\big], \qquad (18)$$

respectively. Here $\dot{u}(2T_s) = u(2T_s) - m(2T_s), \dot{u}(6T_s) = u(6T_s) - m(2T_s), \dot{u}(10T_s) = u(10T_s) - m(2T_s)$ and $\dot{u}(14T_s) = u(14T_s) - m(2T_s)$. Finally, the mean and variance values of the basic non-decimated realization $u(kT_s)\ \forall\, k \in \overline{0, 16}$ could be calculated by the same formulas (1),(2):

$$m(T_s) = \frac{1}{16+1}\left\{\frac{16+2}{2}m(2T_s) + \sum_{l=1}^{8} u\big(T_s(2l-1)\big)\right\}$$
$$= \frac{1}{17}\big[9m(2T_s) + u(T_s) + u(3T_s) + u(5T_s) + u(7T_s) + u(9T_s)$$
$$+ u(11T_s) + u(13T_s) + u(15T_s)\big] \qquad (19)$$

and

$$var(T_s) = \frac{1}{16+1}\left\{\frac{16+2}{2}var(2T_s) + \sum_{l=1}^{8} \dot{u}^2\big(T_s(2l-1)\big)\right\}$$
$$= \frac{1}{17}\big[9var(2T_s) + \dot{u}^2(T_s) + \dot{u}^2(3T_s) + \dot{u}^2(5T_s) + \dot{u}^2(7T_s) + \dot{u}^2(9T_s)$$
$$+ \dot{u}^2(11T_s) + \dot{u}^2(13T_s) + \dot{u}^2(15T_s)\big], \qquad (20)$$

respectively. It should be noted that one can calculate the same mean and variance values by the ordinary formulas according to [2], too. However, recursive calculations according to the formulas (1), (2) allow us to decrease the number of addition operations as compared with expressions given in [2], especially, for enough large N. In such an example, while calculating $m(4T_s), m(2T_s)$ and $m(T_s)$ we avoid three, five, and nine addition operations, respectively. In general, having any realization consisting of 17 samples one needs 28 less operations in comparison with the operations performed using ordinary formulas. On the other hand, there appears one additional multiplication operation in each recursive step.

5. Conclusions

The number of operations for calculating the mean and variance values of decimated signal realizations could be essentially reduced using recursive formulas (1), (2). In such a case, it is no need to store the whole set of decimated realizations in the memory of a computer, only the basic non-decimated realization is required.

References

1. *Å*str*ö*m, On the choice of sampling rates in parametric identification of time series, *Information Sci.*, **1**, 273–278 (1969).
2. J.S. Bendat, A.G. Piersol, *Measurement and Analysis of Random Data*, John Wiley&Sons, inc., N.Y. (1967).
3. B. Liu, F. Mintzer, Calculation of narrow-band spectra by direct decimation, *IEEE Trans. Acoust., Speech, Signal Processing*, **26**, 529–534 (1978).
4. L. Ljung, *System Identification. Theory for the User*, Prentice-Hall PTR, second edition (1999).
5. J.F. MacGregor, Optimal choice of the sampling interval for discrete process control, *Technometrics*, **18**(2), 151–160 (1976).
6. M.P. Quirk, B. Liu, Improving resolution for autoregressive spectral estimation by decimation, *IEEE Trans. Acoust., Speech, Signal Processing*, **31**, 630–637 (1983).
7. R.L. Payne, G.C. Goodwin, M.B. Zarrop, Frequency domain approach for designing sampling rates for system identification, *Automatica*, **11**, 189–191 (1975).
8. J.G. Proakis, D.G. Manolakis, *Digital Signal Processing: Principles Algorithms and Applications*, Prentice-Hall, London (1996).
9. M. Villalba, B.K. Walker, Spectrum manipulation for improved resolution, *IEEE Trans. Acoust., Speech, Signal Processing*, **17**(6), 820–831 (1989).
10. S. Wang, X. Guan, D. Wang, X. Ma and Y. Su, Adaptive selection of sampling interval in frequency domain for estimating the poles of scatterers, *Piersonline*, **3**(1), 106–110 (2007).
11. J. Xin, H. Tsuji, A.Sano, Optimal sampling interval for system identification based on decimation and interpolation, *IEE Proc.-Control Theory Appl.*, **142**(1), 15–22 (1995).

REZIUMĖ

R. Pupeikis. Apie decimuotų sekų statistikų rekurentines išraiškas

Darbe pasiūlytos rekurentinės išraiškos, skirtos decimuotų sekų statistiniams momentams skaičiuoti. Pateiktas pavyzdys 17 ataskaitų realizacijai.