SIGNAL RECOVERY FROM GROUPED DATA

M. Pawlak[1] *U. Stadtmüller*[2]

[1]Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Manitoba, R3T 5V6, Canada
[2]Abteilung für Mathematik III, University of Ulm, D 89069 Ulm, Germany

ABSTRACT

The problem of recovering a signal in the class of band-limited functions is studied. We consider a situation when discrete data points are first grouped to the points of an uniform grid and then the reconstruction is carried out from such a reduced data set. The data grouping is common for computer rounding errors and may also be viewed as a data compression process. The accuracy of the proposed grouping techniques is examined. These results are used to provide an understanding of the number of grid points required to achieve a given level of accuracy.

KEYWORDS: signal reconstruction, band-limited signals, cardinal expansions, data grouping, compression, rate of convergence

1. SIGNAL RECOVERY

The problem of signal reconstruction from sampled data can be found in various signal processing and communication contexts. In most signal processing applications the signal $f(t)$ to be reconstructed from its discrete samples $\{f(j\tau), j = 0, \pm 1, \ldots\}$ is modeled as band-limited. Hence, let $f(t)$ be a band-limited signal with the bandwidth (maximal frequency) Ω. In the sequel, we shall denote such a class of signals by $\mathrm{BL}(\Omega)$. It is well-known, due to the celebrated Whittaker-Shannon sampling theorem, that any $f(t)$ from $\mathrm{BL}(\Omega)$ can be represented as

$$f(t) = \sum_{j=-\infty}^{\infty} f(j\tau)\mathrm{sinc}\left(\frac{\pi}{\tau}(t - j\tau)\right), \tag{1}$$

provided that the sampling rate $0 < \tau \le \pi/\Omega$, see [1] - [3], [8]. Here $\mathrm{sinc}(t) = \sin(t)/t$.

In practice, one observes only a finite number of noisy samples of $f(t)$ and would like to reconstruct $f(t)$ with the best possible degree of accuracy. Hence let the data be generated from the following model

$$y_j = f(j\tau) + \varepsilon_j \quad j = 0, \pm 1, \ldots, \pm n, \tag{2}$$

where $\{\varepsilon_j\}$ is a noise process assumed to be uncorrelated, zero mean and finite variance denoted by σ^2.

[1]Research supported by NSERC A 8131 and Humboldt Foundation

It has been shown in [4] - [7] that one must not replace $f(j\tau)$ in (1) by y_j from (2) in order to obtain converging reconstruction algorithms from noisy data. To alleviate this problem a certain degree of smoothing is required yielding the following reconstruction procedure:

$$f_n(t) = \tau \sum_{|j| \le n} y_j \varphi(t - j\tau), \tag{3}$$

where $\varphi(t) = \sin(\Omega t)/\pi t$ is the reconstruction kernel. See also [5] for a more general class of reconstruction kernels.

We refer to [4] - [7] for detail analysis of statistical properties of $f_n(t)$ and other related signal reconstruction algorithms.

In this paper we examine a new aspect of reconstruction algorithms. Namely, the signal recovery problem under rounding data to grid points defined on an equally spaced mesh. This topic is motivated by the following two important factors:

(a) In many signal processing problems we are frequently faced with large data sets containing thousands samples per second. Therefore standard reconstruction techniques become very time consuming and elaboration of reconstruction methods with a reduced computational complexity is desirable.

(b) Signals are often sampled imprecisely yielding unwanted jitter error. The data grouping approach can transform this nonuniform sampling problem into a standard regular sampling situation.

In the both aforementioned cases the idea of grouping observations into blocks can be applied. Hence, let $c_{-N} < c_{-N+1} < \cdots < c_{N-1} < c_N$ be the equidistant partition with the grid size $h = c_{j+1} - c_j$ and the number of grid points $2N + 1$ smaller than the sample size $2n + 1$. Hence the grid points are on the form $c_k = kh$. A grouping rule may be represented by the weight function $w(t)$ defined on the interval $[-1/2, 1/2)$ such that

$$\hat{y}_k = \sum_{|j| \le n} w((c_k - j\tau)/h) y_j, \quad k = -N, \ldots, N \tag{4}$$

is the weighted sum of the data points $\{y_j, |j| \le n\}$ falling into the k-th block $I_k = \{j : c_k - h/2 < j\tau \le c_k + h/2\}$. Hence, the sample set $\{y_j, |j| \le n\}$ is divided into $2N + 1$ blocks with the block size equals h.

A simple grouping corresponds to the uniform $w(t)$, i.e., $w(t) = \mathbf{1}_{[-1/2,1/2)}(t)$, where $\mathbf{1}_A(t)$ denotes the indicator function of A. In this case the formula in (4) takes the following form $\hat{y}_k = \sum_{j \in I_k} y_j$, i.e., it is a sum of those y_j's which fall into the k-th block I_k. The following weight functions $w(t) = w(1-2|t|)\mathbf{1}_{[-1/2,1/2)}(t)$, $w(t) = 6(1/4 - t^2)\mathbf{1}_{[-1/2,1/2)}(t)$ correspond to the linear and quadratic grouping strategies, respectively.

Data grouping defined in (4) combined with our basic reconstruction scheme in (3) yield the following signal recovery technique:

$$\hat{f}_n(t) = \tau \sum_{|k| \le N} \hat{y}_k \varphi(t - c_k), \tag{5}$$

where $\varphi(t)$ is the reconstruction kernel defined in (3).

2. ACCURACY OF THE RECONSTRUCTION ALGORITHM FROM GROUPED DATA

It is our principal goal in this paper to examine the accuracy of the reconstruction method proposed in (5). Such analysis allows us to determine the compression ratio N/n yielding a given level of accuracy. In fact the estimator $f_n(t)$ requires $O(n)$ evaluations at each time point t. We show that N in (5) can be selected as n^γ, $0 < \gamma \le 1/2$ depending on the behavior of the signal $f(t)$ at $\pm\infty$ and the kind of the weight function $w(t)$ applied. This implies that the number of evaluations (after grouping) in $\hat{f}_n(t)$ is of order $O(n^\gamma)$, $0 < \gamma \le 1/2$.

The following result describes the mean square performance of $\hat{f}_n(t)$. Let us assume that the weight function $w(t)$ is of order $p \ge 2$, i.e., it satisfies the following condition:

$$\int_{-\infty}^{\infty} t^s w(t) dt = \begin{cases} 1 & s = 0 \\ 0 & 0 < s < p \\ \alpha \ne 0 & s = p \end{cases}. \tag{6}$$

Theorem 1 *Let $f \in BL(\Omega)$ and let*

$$|f(t)| \le a|t|^{-(r+1)}, |t| > 0 \quad \text{with} \quad 0 \le r < \infty. \tag{7}$$

Suppose that (6) is met. If $\tau < h \le \pi/\Omega$, then for $|t| < Nh$ we have

$$E(\hat{f}_n(t) - f(t))^2 \le 4\frac{\sigma^2\Omega}{\pi}\tau + \eta_1(Nh)^{-2(r+1)} + \eta_2 h^{2p} + \eta_3\tau^2,$$

where η_1, η_2, η_3 are some positive constants.

It is worth noting that the parameter r appearing in (7) controls the behavior of $f(t)$ at $\pm\infty$.

It has been shown in [5] that the mean square error of the reconstruction technique in (3) cannot tend faster to zero than $O(n^{-2(r+1)/(2r+3)})$. The result of Theorem 1 reveals that the same rate can be reached by the reduced complexity reconstruction technique in (5) if one chooses N of order $O(n^{(p+r+1)/p(2r+3)})$. Hence for r varying from ∞ to 0 the computational complexity of (5) is $O(n^\lambda)$, $1/2p < \lambda \le (p+1)/3p$. For the simplest grouping strategy with $w(t) = \mathbf{1}_{[-1/2,1/2)}(t)$ the number of grid points can be selected as $N = \lfloor n^{(r+3)/2(2r+3)} \rfloor$, where $\lfloor a \rfloor$ is an integer part of a. The latter resolves to $N = \lfloor n^{1/2} \rfloor$ for the slowest decaying (as $1/|t|$) band-limited signals.

The aforementioned results can also be extended to the case of irregular sampling, i.e., when one instead of (2) has $y_j = f(t_j) + \varepsilon_j$, for a certain sequence of points $\{t_j\}$.

REFERENCES

[1] J.R. Higgins, *Sampling Theory in Fourier and Signal Analysis*, Clarendon Press, Oxford, 1996.

[2] R.J. Marks II, *Introduction to Shannon Sampling and Interpolation Theory*, Springer-Verlag, New York, 1991.

[3] R.J. Marks II, (ed.), *Advances Topics in Shannon Sampling and Interpolation Theory*, Springer-Verlag, New York, 1993.

[4] M. Pawlak and E. Rafajlowicz, "On restoring band-limited signals," *IEEE Trans. Information Theory*, vol. 40, pp. 1490-1503, 1994.

[5] M. Pawlak and U. Stadtmüller, "Recovering band-limited functions under noise," *IEEE Trans. Information Theory*, vol. 44, pp. 538-545, 1996.

[6] M. Pawlak, E. Rafajlowicz and A. Krzyzak, "Exponential weighting algorithms for reconstruction of bandlimited signals," *IEEE Trans. Signal Processing*, vol. 44, pp. 538-545, 1996.

[7] M. Pawlak and U. Stadtmüller, "Kernel regression estimators for signal recovery," *Statistics and Probability Letters*, 1996, to appear.

[8] A.I. Zayed, *Advances in Shannon's Sampling Theory*, CRC Press, Boca Raton, 1993.