Research Article

Recovery of Missing Samples with Sparse Approximations

Benjamin G. Salomon

School of Engineering, Ruppin Academic Center, 40250 Emek Hefer, Israel

Correspondence should be addressed to Benjamin G. Salomon; bennysal123@gmail.com

In most missing samples problems, the signals are assumed to be bandlimited. That is, the signals are assumed to be sparsely approximated by a known subset of the discrete Fourier transform basis vectors. We discuss the recovery of missing samples when the signals can be sparsely approximated by an unknown subset of certain unitary basis vectors. We propose the use of the orthogonal matching pursuit to recover missing samples by sparse approximations.

1. Introduction

Discrete signals are usually represented by their samples taken on a uniform sampling grid. However, in many applications, it may happen that some samples are lost or unavailable. In such cases, it is required to convert the irregularly sampled signal to a regularly sampled one, that is, to restore the missing samples. One approach for the recovery of missing data in discrete signals is based on the assumption that the underlying continuous-time signals are bandlimited. The celebrated sampling theorem by Whittaker [1], Kotel'nikov [2], and Shannon [3] implies that any continuous-time bandlimited signal can be reconstructed by its regularly spaced samples if the sampling frequency is higher than two times the maximum frequency component of the signal. The solution of the nonuniform sampling problem poses more difficulties and there exists a vast literature dealing with necessary and sufficient conditions for unique reconstruction and methods for reconstructing a function from its samples, for example, [4–6]. The numerical reconstruction methods, however, have to operate in a finite-dimensional model, whereas the theoretical results are usually derived for the continuous-time bandlimited functions (an infinite-dimensional subspace). The use of a truncated sampling series results with a finite-dimensional model but may lead, however, to severely ill-posed numerical problems [7].

Another approach is to address this problem in a finite-dimensional model of discrete bandlimited signals. A discrete bandlimited (DBL) signal has a sparse representation in terms of a certain unitary basis vector (e.g., discrete Fourier transform). That is, the signal can be represented by only a known subset of the unitary basis vectors. The recovery of missing samples is, in this case, equivalent to solving a linear system of equations [8]. In this paper, we focus on the recovery problem, when the a priori knowledge is that the signal is sparsely represented by an *unknown* subset of certain unitary basis vectors. This problem is much harder to solve, and there is an infinite number of possible solutions. Our approach is to choose the sparsest solution. That is, we are interested in approximating the known samples with a minimum number of basis vectors. Standard methods for solving linear systems of equations cannot provide the sparsest solution. We suggest the use of the orthogonal matching pursuit algorithm [9–11] for determining the sparsest approximation to the given samples from which the unknown samples can be determined.

2. Recovery of Missing Samples of Discrete Bandlimited Signals

Let $x \in \mathbb{R}^N$ be a discrete signal with N samples. That is, x can be described by an N-dimensional real vector whose elements are denoted by $x[0], x[1], \ldots, x[N-1]$. These elements correspond to the samples of the signal. Let Φ be an $N \times N$ orthonormal transform matrix; that is, $\Phi^T \Phi = \Phi \Phi^T = I$, where I is the $N \times N$ identity matrix and Φ^T

is the transpose of Φ. The orthonormal matrix Φ defines an orthonormal transform of the vector x, denoted by $\widehat{x} \in \mathbb{R}^N$, which is by definition the signal $\widehat{x} = \Phi x$. The inverse transform of $\widehat{x}$ is by definition the signal $x = \Phi^T \widehat{x}$. In the complex case, the vectors belong to the space $\mathbb{C}^N$ (space of N-dimensional complex vectors), and the conjugate transpose of the unitary matrix Φ has to be taken when computing the inverse transform.

The columns of Φ^T define an orthonormal basis of $\mathbb{R}^N$. That is, each signal $x \in \mathbb{R}^N$ can be represented by a linear combination of the columns of Φ^T, where the coefficients in this linear combination are given by the transform coefficients $\widehat{x}$. Let $(\{\phi_r\}_{r=0,1,\dots,N-1})$ be the rows of Φ^T (i.e., the columns of the matrix Φ). It follows that for each $x \in \mathbb{R}^N$

$$x[k] = \sum_{m=0}^{N-1} \widehat{x}[m]\, \phi_k[m], \quad k = 0, 1, \dots, N-1. \tag{1}$$

Let Q be a K-size proper subset of $\{0, 1, \dots, N-1\}$ and assume that only the $K < N$ samples $\{x[k]\}$, $k \in Q$, are known. Consequently, the available K signal samples define systems of K equations for the N transform coefficients:

$$x[k] = \sum_{m=0}^{N-1} \widehat{x}[m]\, \phi_k[m], \quad k \in Q. \tag{2}$$

The recovery of missing samples from an arbitrary signal is of course impossible: there is an infinite number of solutions to the underdetermined system of equations (2). The signals that occur in applications, however, often satisfy known constraints. The constraint used in this paper is called discrete band-limitedness: the signal can be represented by only a known subset of the columns of Φ^T. That is, the signal is a linear combination of certain subset of certain orthonormal basis vectors of $\mathbb{R}^N$.

Let R be a P-size proper subset of $\{0, 1, \dots, N-1\}$, where $P \leq K$. A discrete bandlimited signal approximation to the signal x is obtained by

$$x^{\mathrm{BL}}[k] = \sum_{m \in R} \widehat{x}[m]\, \phi_k[m], \quad k = 0, 1, \dots, N-1. \tag{3}$$

That is, we assume that all transform coefficients with indices $r \notin R$ are equal to zero.

For example, let Φ be the $N \times N$ discrete Fourier transform (DFT) matrix. That is, Φ is the $N \times N$ unitary matrix with elements (j is the imaginary unit)

$$\Phi_{kl} = \frac{1}{\sqrt{N}} e^{-j(2\pi/N)kl}. \tag{4}$$

A signal x is bandlimited if its DFT $\widehat{x} = \Phi x$ vanishes on some fixed nonempty proper subset S of $\{0, 1, \dots, N-1\}$. The complement of S is the subset R. When the subset R is the set of $2M + 1$ elements (i.e., $P = 2M + 1$)

$$\{0\} \cup \{1, 2, \dots, M\} \cup \{N - M, \dots, N - 1\}, \tag{5}$$

or equivalently (by periodicity)

$$\{0, \pm 1, \pm 2, \dots, \pm M\}, \tag{6}$$

the signal is called a low-pass signal in DFT domain.

Another example is the discrete cosine transform. Let Φ be the $N \times N$ discrete cosine transform (DCT) matrix. That is, Φ is the $N \times N$ orthonormal matrix with elements

$$\Phi_{kl} = \begin{cases} \dfrac{1}{\sqrt{N}}, & k = 0, 0 \leq l \leq N - 1, \\[3mm] \sqrt{\dfrac{2}{N}} \cos \dfrac{\pi(2l+1)k}{2N}, & 1 < k \leq N - 1, \\ & 0 \leq l \leq N - 1. \end{cases} \tag{7}$$

A signal x is bandlimited if its DCT $\widehat{x} = \Phi x$ vanishes on some fixed nonempty proper subset S of $\{0, 1, \dots, N-1\}$. When the complement of S (i.e., subset R) is the set of P elements

$$\{0, 1, 2, \dots, P - 1\}, \tag{8}$$

the signal is called a low-pass signal in DCT domain.

Assuming that x is a discrete bandlimited signal (i.e., $x[k] = x^{\mathrm{BL}}[k]$, $k = 0, \dots, N-1$), then, the available signal samples, $x[k]$, $k \in Q$, can now be expressed as

$$x[k] = \sum_{m \in R} \widehat{x}[m]\, \phi_k[m], \quad k \in Q. \tag{9}$$

Equation (9) is a linear system consisting of K equations with $P \leq K$ unknowns. The existence and uniqueness of the solution depend, for any transform, on the subsets Q and R (see, e.g., [8] for a discussion on the recovery of samples when the signals are bandlimited in the discrete Fourier transform domain). When the system of equations (9) has a full column rank and $K = P$, we can determine the unique exact solution. When the system has a full column rank and $K > P$ (overdetermined system of equations), we will be interested in the least squares solution.

3. Sparse Approximations

In the previous section, we discussed the recovery of missing samples, when the band-limitedness was known a priori. That is, we had the a priori knowledge of which transform coefficients are equal to zero. The recovery of the missing samples is, in this case, equivalent to solving a linear system of equations for which many algorithms can be used. If we know that the signal is a low-pass bandlimited signal (e.g., in DFT or DCT domain), but the bandwidth is not given, we may start with a low bandwidth model and solve the resulting linear system of equations for increasing bandwidths until a sufficient accuracy in approximating the given samples is obtained. If the only a priori knowledge is that the signal is sparse in terms of the basis vectors, that is, the signal can be represented (or accurately approximated) with a few *unknown* basis vectors, the situation is very different. In this case, we have also to determine the basis vectors that sparsely approximate the signal.

We have to solve the underdetermined system of equations (2):

$$x(k) = \sum_{m=0}^{N-1} \widehat{x}(m)\, \phi_k(m), \quad k \in Q, \tag{10}$$

where Q is the K-size subset of indices of known samples. The underdetermined system of equations (2) has an infinite number of solutions. Since we know that the signal is sparsely approximated by the basis vectors, one reasonable approach is to choose the sparsest solution among the infinite number of solutions. That is, the optimal approximation is defined as either the sparsest approximation (i.e,. with the fewest basis vectors) that yields an approximation error that is smaller than a prespecified threshold or the approximation using a fixed number of basis vectors that minimizes the approximation error. Finding these approximations is an NP hard problem [12, 13].

We will use the following notations. Let $f \in \mathbb{R}^K$ be a signal of K samples that contains the given samples; that is, $f[n] = x[k_n]$, $k_n \in Q$, $n = 0, 1, \ldots, K-1$. Let φ_r, $r = 0, 1, \ldots, N-1$, be the columns of the matrix Φ^T. For each $\gamma = 0, 1, \ldots, N-1$, let $g_\gamma = h_\gamma / \|h_\gamma\|$ be a unit-norm vector of K samples, where $h_\gamma[n] = \varphi_\gamma[k_n]$, $k_n \in Q, n = 0, 1, \ldots, K-1$, and the norm is the standard l_2 norm in $\mathbb{R}^K$: $\|f\|^2 = \sum_{n=0}^{K-1} f^2[n]$. The set of N unit-norm vectors g_γ will be called a dictionary. The problem is to determine a sparse approximation to the vector f with the dictionary vectors g_γ, $\gamma = 0, 1, \ldots, N-1$.

After determining a sparse solution using L dictionary vectors:

$$f \approx \sum_{i=1}^{L} \alpha_{\gamma_i} g_{\gamma_i} = \sum_{i=1}^{L} \beta_{\gamma_i} h_{\gamma_i}, \qquad (11)$$

where

$$\beta_{\gamma_i} = \frac{\alpha_{\gamma_i}}{\|h_{\gamma_i}\|}, \qquad (12)$$

the original signal x will approximated by

$$x \approx \sum_{i=1}^{L} \beta_{\gamma_i} \varphi_{\gamma_i}, \qquad (13)$$

from which the missing samples can be determined.

Several algorithms have been proposed for reducing the computational complexity by searching for efficient but nonoptimal approximations. The matching pursuit (MP) [10, 11] is a popular iterative greedy algorithm for approximate decomposition that addresses the sparsity issue directly. Vectors are selected one by one from the dictionary, picking at each iteration the vector that best correlates with the present residual, and thus optimizing the signal approximation (in terms of energy) at each iteration. An intrinsic feature of this algorithm is that when stopped after a few steps, it yields an approximation using only a few dictionary vectors.

We consider the space $\mathbb{R}^K$ of real-valued signals of size K. Let $D = \{g_\gamma\}$, $\gamma = 0, 1, \ldots, N-1$, be a dictionary of $N > K$ vectors, having a unit norm. Let $R_n f$ be the residual of an n term approximation of a given signal $f \in \mathbb{R}^K$. MP subdecomposes the residue $R_n f$ by projecting it on the dictionary vector that matches $R_n f$ best. Starting from an initial approximation $f_0 = 0$ and residue $R_0 f = f$, the MP algorithm builds up a sequence of sparse approximations

stepwise. MP begins by projecting f on a vector $g_{\gamma_0} \in D$ and computing the residue $R_1 f$:

$$f = \langle f, g_{\gamma_0} \rangle g_{\gamma_0} + R_1 f, \qquad (14)$$

where $R_1 f$ is the residual vector after best approximating (l_2 norm sense) $R_0 f = f$ with g_{γ_0}. Since $R_1 f$ is orthogonal to g_{γ_0},

$$\|f\|^2 = \left| \langle f, g_{\gamma_0} \rangle \right|^2 + \|R_1 f\|^2. \qquad (15)$$

It follows that to minimize $\|R_1 f\|$, we must choose $g_{\gamma_0} \in D$ such that $|\langle f, g_{\gamma_0} \rangle|$ is maximum. The process is iterated on the residual. Suppose that $R_n f$ for $n \geq 0$ is already computed. The next iteration chooses g_{γ_n} such that

$$g_{\gamma_n} = \arg \min_{g_\gamma \in D} \|R_{n+1} f\|^2 = \arg \min_{g_\gamma \in D} \|R_n f - a_n g_\gamma\|^2. \qquad (16)$$

For unit-norm vectors, the last condition is equivalent to

$$g_{\gamma_n} = \arg \max_{g_\gamma \in D} \left\| \langle g_\gamma, R_n f \rangle \right\|. \qquad (17)$$

That is, the optimal vector is the one which best correlates with the residual. The MP projects $R_n f$ on the chosen vector:

$$R_n f = \langle R_n f, g_{\gamma_n} \rangle g_{\gamma_n} + R_{n+1} f. \qquad (18)$$

The orthogonality of $R_{n+1} f$ and g_{γ_n} implies

$$\|R_n f\|^2 = \left| \langle R_n f, g_{\gamma_n} \rangle \right|^2 + \|R_{n+1} f\|^2. \qquad (19)$$

Summing (18) from n between 0 and $L-1$ (for any integer $L > 0$) yields

$$f = \sum_{n=0}^{L-1} \langle R_n f, g_{\gamma_n} \rangle g_{\gamma_n} + R_L f. \qquad (20)$$

Similarly, summation of (19) from n between 0 and $L-1$ gives

$$\|f\|^2 = \sum_{n=0}^{L-1} \left\| \langle R_n f, g_{\gamma_n} \rangle \right\|^2 + \|R_L f\|^2. \qquad (21)$$

When dealing with finite-dimension signals (as in our case), $\|R_L f\|$ converges exponentially to 0 when L tends to infinity [10].

The approximations of MP are improved by orthogonalizing the directions of projection with a Gram-Schmidt procedure [9]. The resulting orthogonal MP (OMP) converges with a finite number of steps, which is not the case for a standard MP. The vector g_{γ_n} selected by the MP is a priori not orthogonal to the previously selected vectors g_{γ_p}, $p = 0, 1, \ldots, n-1$. After subtracting the projection of $R_n f$ over g_{γ_n}, new components are introduced in the direction of g_{γ_p}, $p = 0, 1, \ldots, n-1$. This is avoided by projecting the residues on an orthonormal set of vectors u_{γ_p}, $p = 0, 1, \ldots, n-1$, that span the same subspace that is spanned by g_{γ_p}, $p = 0, 1, \ldots, n-1$. That is, in each iteration, we determine the
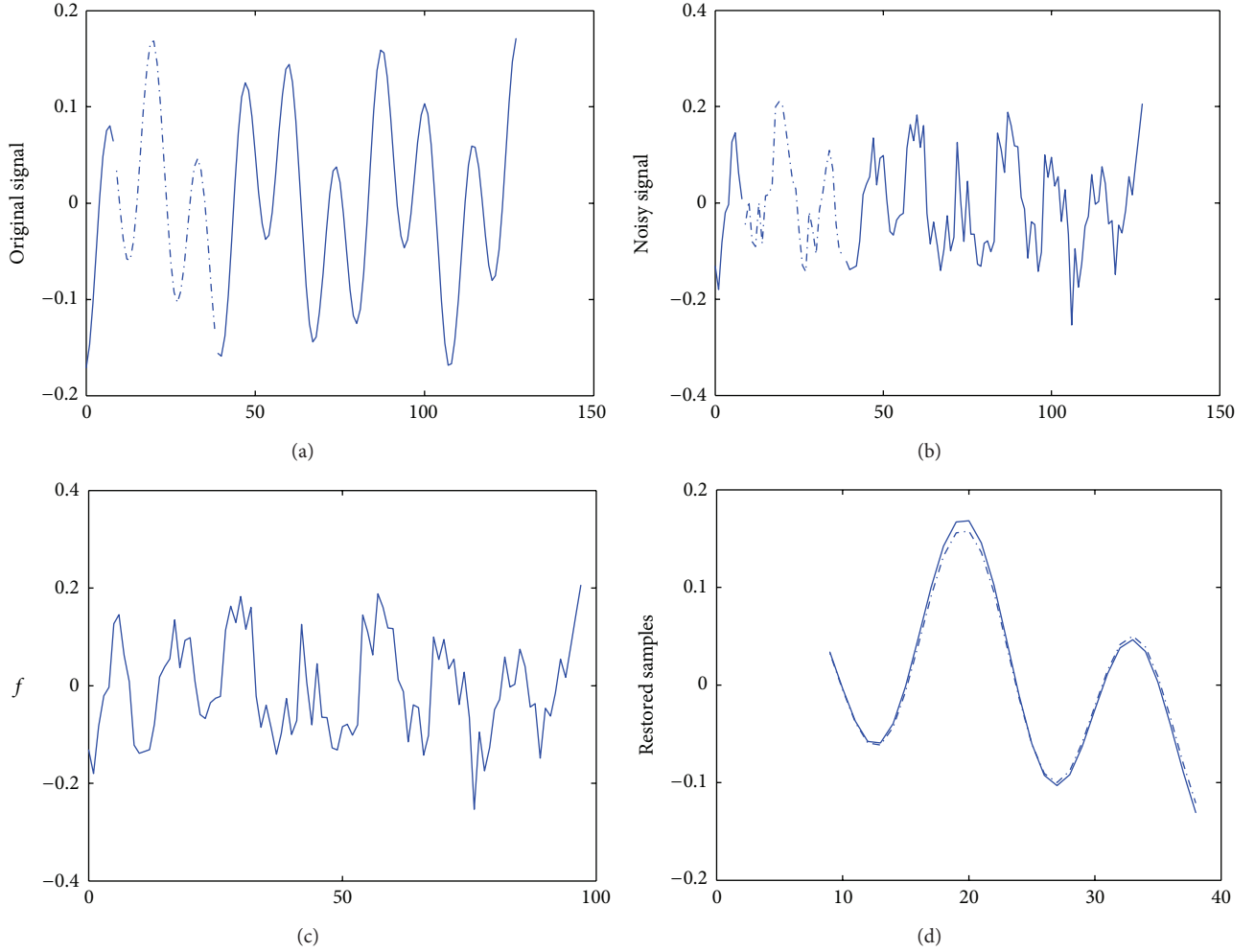
FIGURE 1: (a) Original signal (dash-dot: missing samples), (b) noisy signal (dash-dot: missing samples), (c) the vector f containing the noisy given samples, and (d) original missing samples and the restored samples (dash-dot).

best approximation of the residual in the subspace spanned by u_{γ_p}, $p = 0, 1, \ldots, n-1$. When an approximation of sufficient accuracy is obtained, to expand f over the original dictionary vectors g_{γ_p}, $p = 0, 1, \ldots, n-1$, we perform a change of basis by expanding u_m in g_{γ_p}, $p = 0, 1, \ldots, n-1$.

We demonstrate the applicability of the approach by example. The original 128 samples length signal is the linear combination of two basis vectors of the DCT with randomly chosen coefficients: $x = -0.5465\varphi_7 - 0.8468\varphi_{19}$. The given signal is a noisy version of the original signal with additive white Gaussian noise with standard deviation equal to 0.05. We assume that 30 contiguous samples (9 to 38) of the noisy signal are missing. The results of the OMP are depicted in Figure 1, and we can see that the missing samples were approximately restored. The l_2 norm of the missing samples is 0.4339, while the norm of the error is 0.0315. If the samples of the original exact DBL signal x are given, the missing 30 samples are completely recovered by the OMP procedure.

4. Conclusions

In this paper, we addressed the recovery of missing samples of discrete bandlimited signals. We have suggested the use of the orthogonal matching pursuit to recover missing samples by sparse approximations, when the signal can be sparsely represented by an unknown subset of basis vectors of a certain orthonormal transform.

References

[1] E. T. Whittaker, "On the functions which are represneted by the expansion of the interpolation theory," in *Proceedings of the Royal Society of Edinburgh*, vol. 35, pp. 181–194, 1915.

[2] V. A. Kotel'nikov, "On the carrying capacity of "Ether" and wire in telecommunications," in *Material for the First All-Union Conference on Questions of Communications*, Svyazi RKKA, Moscow, Russia, 1963.

[3] C. E. Shannon, "Communication in the presence of noise," in *Proceedings of the Institution of Radio Engineers*, vol. 37, no. 1, pp. 10–21, 1949.

[4] H. G. Feichtinger and K. H. Grochenig, "Theory and practice of irregular sampling," in *Wavelets: Mathematics and Applications*, J. Benedetto and M. Frazier, Eds., vol. 1994, pp. 305–363, CRC Press, Boca Raton, Fla, USA.

[5] H. J. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Mathematica*, vol. 117, no. 1, pp. 37–52, 1967.

[6] J. L. Yen, "On nonuniform sampling of bandwidth-limited signals," *IRE Transactions on Circuit Theory*, vol. 3, no. 4, pp. 251–257, 1956.

[7] T. Strohmer, "Numerical analysis of the non-uniform sampling problem," *Journal of Computational and Applied Mathematics*, vol. 122, no. 1, pp. 297–316, 2000.

[8] P. J. S. G. Ferreira, "Iterative and noniterative recovery of missing samples for 1-D band-limited signals," in *Nonuniform Sampling: Theory and Practice*, F. Marvasti, Ed., pp. 235–278, Kluwer Academic/Plenum, New York, NY, USA, 2001.

[9] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Proceedings of the 27th Asilomar Conference on Signals, Systems & Computers*, pp. 40–44, Pacific Grove, Calif, USA, November 1993.

[10] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

[11] S. Mallat, *A Wavelet Tour of Signal Processing, 2nd Edition*, Academic Press, New York, NY, USA, 1999.

[12] G. Davis, S. Mallat, and M. Avellaneda, "Adaptive greedy approximations," *Constructive Approximation*, vol. 13, no. 1, pp. 57–98, 1997.

[13] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM Journal on Computing*, vol. 24, no. 2, pp. 227–234, 1995.