Graphical Abstract

Slepians.jl: A Julia package to generate optimal functions which are simultaneously constrained in space and spectral domains

Charlotte L. Haley, M. Anitescu, V. Rao, M. Krogstad, R. Osborn, S. Rosenkranz

Highlights

Slepians.jl: A Julia package to generate optimal functions which are simultaneously constrained in space and spectral domains

Charlotte L. Haley, M. Anitescu, V. Rao, M. Krogstad, R. Osborn, S. Rosenkranz

- Research highlight 1
- Research highlight 2

Slepians.jl: A Julia package to generate optimal functions which are simultaneously constrained in space and spectral domains

Charlotte L. Haley[a], M. Anitescu[a], V. Rao[a], M. Krogstad[b], R. Osborn[c], S. Rosenkranz[c]

[a] *Argonne National Laboratory, Division of Mathematics and Computer Science, Lemont, 60439, IL, USA*
[b] *Argonne National Laboratory, Advanced Photon Source, Lemont, 60439, IL, USA*
[c] *Argonne National Laboratory, Materials Science Division, Lemont, 60439, IL, USA*

Abstract

In a series of six landmark papers of Slepian, Landau, and Pollak, a set of functions are described which solve what is known as the "concentration problem". The concentration problem refers to finding the subspace of functions having limited support in, say, time, which simultaneously have a Fourier transform for which the bulk of its nonzero mass in a small constrained interval. While in some special discrete cases, it is possible to solve a simple, possibly tridiagonal or Toeplitz eigenvalue problem, in others it is necessary to use numerical integration to solve the defining Fredholm integral equation of the second kind. This package produces function values where the problem dimension is up to three. We rely on the work of Simons, Wang, [1] and others, for the two dimensional case but to our knowledge the three dimensional case is novel to this package. In addition, we implement the missing data Slepian sequences of Chave [2] as well as the generalized Slepian sequences of Bronez [3] which solve the 2D concentration problem on an unequally spaced spatial grid. We give a novel real data application of these functions to the analysis of the pair distribution function, which summarizes the probability of the relative positions of pairs of atoms in a crystal lattice, and is the Fourier transform of the single crystal x-ray scattering (which is understood to be like the spectrum of an unobserved process).

The companion package to this work represents enhanced computational efficiency above and beyond that which is available at present, and can solve problems in 3D, which is entirely novel.

Preprint submitted to Computer Physics Communications *June 20, 2023*

Keywords:

1. Introduction

In the early 1960s, the prolate spheroidal wave functions (pswf) of Slepian, Landau, and Pollak [4, 5, 6, 7] were a series of landmark papers published in the Bell System Technical Journal which described a special one-dimensional class of functions having finite extent in time but constrained extent in the spectral (or frequency) domain.

One of the earliest applied contributions making use of these functions was the multitaper method of Thomson [8, 9] which, simply put, used the discrete analog of the pswfs, [10], as data windows to constrain spectral leakage and reduce variance when estimating quantities in the spectral domain such as the power spectrum, coherency, and so forth. While the multitaper estimator introduces a moderate amount of localized bias, the reduction in far-field bias and variance can be shown to be large, so much so that in comparsion with Welch spectrum estimators, when two of bias, variance, and bandwidth are kept constant, it can be shown [11] that the third quantity is always smaller when one employs the multitaper spectrum estimator.

The pwsfs $g(x)$ solve the concentration problem in which a bandlimited signal is optimized in a mean squared sense to live strictly within a predefined interval in time [12]. More generally, one can refer to the problem of constraining the spacelimited function defined over a particular domain of interest, where the space of origin may be Cartesian [13] on the sphere [14] or other.

2. Mathematical Preliminaries

The concentration problem of Slepian, Landau and Pollak, seeks to find bandlimited functions $g(t)$, i.e. functions satisfying

$$g(t) = \frac{1}{2\pi} \int_{-W}^{W} G(\omega) e^{i\omega t} d\omega \tag{1}$$

where $G(\omega)$ denotes the Fourier transform of the function $g(t)$, i.e.

$$G(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \tag{2}$$

that maximize the ratio

$$\lambda = \frac{\int_{-T}^{T} g^2(t)dt}{\int_{-\infty}^{\infty} g^2(t)dt}. \tag{3}$$

λ is referred to as the concentration. It was shown that these functions satisfy the Fredholm integral equation of the second kind

$$\int_{-T}^{T} \frac{\sin[W(t-s)]}{\pi(t-s)} g(s)ds = \lambda g(t) \tag{4}$$

for real t.

3. Discrete prolate spheroidal sequences

The discrete prolate spheroidal sequences [10], $v_n^{(k)}(N, W)$, are the discrete analog to the above functions $g(t)$. As such, they are the eigenvector solutions to the eigenvalue problem

$$\sum_{m=0}^{N-1} \frac{\sin 2\pi W(n-m)}{\pi(n-m)} v_m^{(k)}(N, W) = \lambda_k(N, W) \cdot v_n^{(k)}(N, W) \tag{5}$$

where N is the length of the sequence, $n = 0, \ldots, N-1$, W is the bandwidth, K is the number of tapers, and $\lambda_k(N, W)$ is the eigenvalue corresponding to the kth taper. Fig. 1 shows the discrete prolate spheroidal sequences calculated using $N = 1,024$, $NW = 4$, and $K = 8$.

The dpss's solve the problem of concentrating, in frequency, the most amount of mass under the interval $(-W, W)$, while having only finite extent in time. The first $2NW$ tapers have eigenvalues close to one, whereby their magnitude drops off rapidly thereafter.

Of note is that the dpss's are never computed using the above equation because of numerical instability. Grunbaum [15] showed that the associated differential operator commutes with a symmetric tridiagonal matrix T, which is

$$T_{nn} = [(N-1-2n)/2]^2 \cos(2\pi W), \quad n = 0, \ldots, N-1 \tag{6}$$

on the diagonal and

$$T_{n,n+1} = T_{n,n-1} = (n+1)(N-n-1)/2, \quad n = 0, \ldots, N-2 \tag{7}$$

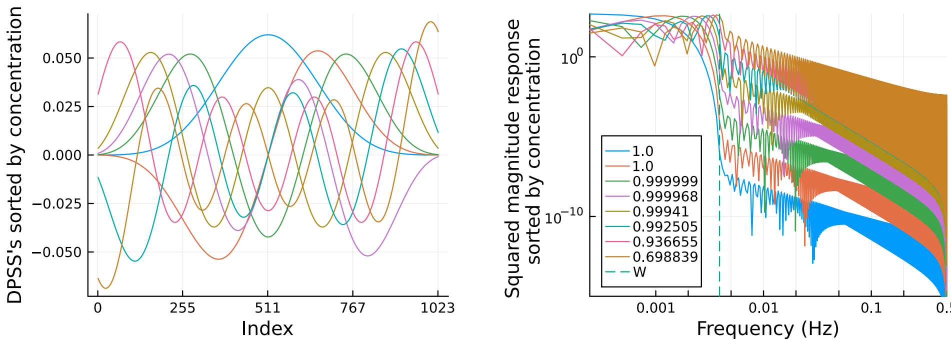

Figure 1: Standard discrete prolate spheroidal sequences computed using Eqn. (5) and parameters $N = 1,024$, $NW = 4$, and $K = 8$. Left panel shows the sequences in the time domain, while the right panel shows the squared magnitude of their Fourier transforms. The sequences have their mass (indicated by the value of λ in the legend – the first two values of λ are $0.9999999997055463, 0.999999972328674, 0.9999987902597943$, which have been rounded in the legend) concentrated in the region $(-W, W)$, where W is shown by the vertical line in the right panel.

on the sub and super diagonals. The matrix T has the same eigenvectors as the problem above and the eigenvalues can be found by substitution into the original equation. The tridiagonal formulation reduces the computation to one which is soluble in $\mathrm{O}(N)$ time, and the solution tends to be more accurate.

Generate a Durrani-Chapman filter [16]

4. Generalized Prolate Spheroidal Sequences

The missing-data multitaper method requires its own set of optimal tapers [8, 3, 2], the *missing-data prolate spheroidal sequences* or missing-data Slepian sequences. These sequences extend the notion of maximally bandlimited orthogonal sequences, [10], to those sampled on a grid where there is missing-data. Define the bandwidth W for the desired spectral window, and let the sequences be given on the length N index set $\{t_n\}_{n=0}^{N-1}$ with unit sampling except where there are gaps.

The missing-data Slepian sequences solve the eigenvalue problem, equations (21)-(23) in [17],

$$\lambda_k v_n^k = \sum_{m=0}^{N-1} \frac{\sin 2\pi W(t_n - t_m)}{\pi(t_n - t_m)} v_m^k. \tag{8}$$

The kth missing-data Slepian sequence is denoted v_n^k at time t_n, and and λ_k is its associated eigenvalue, sorted in decreasing order $1 > \lambda_0 > \lambda_1 > \ldots > \lambda_{N-1} > 0$. Note that both of these explicitly depend on N and W, but for simplicity this is suppressed in the notation. When there are no missing values, the sinc-like matrix in (8) is Toeplitz and has a special form which allows for numerically effective eigenvalue routines.

By applying the nonuniform discrete Fourier transform to the generalized data taper in (8), one obtains the missing-data prolate spheroidal functions, or *missing-data Slepian functions*

$$V^k(f) = \sum_{n=0}^{N-1} v_n^k e^{-2\pi i t_n f}. \tag{9}$$

The missing-data Slepian sequences form an orthonormal set on the grid $\{t_n\}_{n=0}^{N-1}$; and the missing-data Slepian functions are orthonormal on $(-1/2, 1/2)$ and orthogonal on $(-W, W)$.

Show a figure as an example

5. Higher dimensional Cartesian Slepian functions

While what follows is contained entirely in [13], we introduce the mathematical preliminaries and establish notation in this section.

In two dimensions, one can consider the problem of concentrating a function in a region of physical space for application to data collected from the natural sciences. For this application we are concerned with the numerical solution to a Fredholm integral equation of the second kind.

6. Examples

6.1. One Dimensional Examples

6.1.1. Missing data spectral analysis

6.2. X-ray Scattering of single crystals

We give a novel real data application of these functions to the analysis of the pair distribution function, which summarizes the probability of the relative positions of pairs of atoms in a crystal lattice, and is the Fourier transform of the single crystal x-ray scattering (which is understood to be like the spectrum of an unobserved process).

7. Software

'Slepians.jl' is a Julia package for producing simultaneously time and bandlimited sequences, as well as space and spectrally limited functions up to three dimensions. These functions were introduced by Slepian, Landau, and Pollak in the 1960s and have had particular application in diverse fields, perhaps most notably as foundational for the multitaper method for spectrum analysis [8] which extends both to 2D Cartesian domains [1] as well as the sphere [14].

The high-level character of Julia allows for widely readable and extendible codes, while the low-level functionality provides speed and efficiency. The 'Multitaper.jl' package provides a user-friendly implementation of many of the basic concepts. Implementations of higher-dimensional Slepian tapers on Cartesian domains [1, 18]. In addition, we provide tutorial-style notebooks to allow accessibility to those new to these concepts or to Julia in general.

'Slepians.jl' has been used in the context of 3D pair distribution function analysis of crystal structure.

8. Other software

The Matlab software of Frederik Simons's research group is available on github under `https://github.com/csdms-contrib`, of particular note is [?] which relies on [?] and [?]. These packages

9. To contribute

We welcome input of any kind via bitbucket issues or by pull requests. Support requests can kindly be directed to haley@anl.gov.

10. Acknowledgements

We acknowledge contributions from David J. Thomson, Sally Dodson-Robinson during the writing of these codes. We also gratefully acknowledge the help of our reviewers in editing the code repository.

This work was supported by the U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research, under contract number DE-AC02-06CH11357.

Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan. http://energy.gov/downloads/doe-public-access-plan

References

[1] F. J. Simons, D. V. Wang, Spatiospectral concentration in the Cartesian plane, GEM-International Journal on Geomathematics 2 (1) (2011) 1–36. doi:10.1007/s13137-011-0016-z.

[2] A. D. Chave, A multitaper spectral estimator for time-series with missing data, Geophysical Journal International 218 (3) (2019) 2165–2178. doi:10.1093/gji/ggz280.

[3] T. P. Bronez, Spectral estimation of irregularly sampled multidimensional processes by generalized prolate spheroidal sequences, IEEE Trans. on Signal Processing 36 (1988) 862–873.

[4] D. Slepian, H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty -I, Bell System Tech. J. 40 (1) (1961) 43–64.

[5] H. J. Landau, H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty-II, Bell System Tech. J. 40 (1961) 65–84.

[6] H. J. Landau, H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty-III, Bell System Tech. J. 40 (1962) 1295–1336.

[7] D. Slepian, Prolate spheroidal wave functions, Fourier analysis and uncertainty -IV, Bell System Tech. J. 43 (1964) 3009–3057.

[8] D. J. Thomson, Spectrum estimation and harmonic analysis, Proceedings of the IEEE 70 (9) (1982) 1055–1096. doi:10.1109/proc.1982.12433.

[9] D. B. Percival, A. T. Walden, Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques, Cambridge Univ. Press, Cambridge, 1993.

[10] D. Slepian, Prolate spheroidal wave functions, Fourier analysis, and uncertainty V: The discrete case, Bell System Tech. J. 57 (5) (1978) 1371–1429.

[11] T. P. Bronez, On the performance advantage of multitaper spectral analysis, IEEE Trans. on Signal Processing 40 (1992) 2941–2946.

[12] D. Slepian, Some comments on Fourier analysis, uncertainty, and modeling, SIAM Rev. 25 (1983) 379–393.

[13] F. J. Simons, D. V. Wang, Spatiospectral concentration in the cartesian plane, International Journal on Geomathematics 2 (1) (2011) 1–36.

[14] F. J. Simons, F. A. Dahlen, M. A. Wieczorek, Spatiospectral concentration on a sphere, SIAM review 48 (3) (2006) 504–536.

[15] F. A. Grünbaum, Toeplitz matrices commuting with tridiagonal matrices, Linear Algebra and Its App. 40 (1981) 25–36.

[16] T. S. Durrani, R. Chapman, Optimal all-pole filter design based on discrete prolate spheroidal sequences, IEEE Trans. on Signal Processing 32 (1984) 716–721.

[17] A. D. Chave, D. S. Luther, D. J. Thomson, High-Q spectral peaks and nonstationarity in the deep ocean infragravity wave band: Tidal harmonics and solar normal modes, Journal of Geophysical Research: Oceans (2019).

[18] C. J. Geoga, C. L. Haley, A. R. Siegel, M. Anitescu, Frequency–wavenumber spectral analysis of spatio-temporal flows, Journal of Fluid Mechanics 848 (2018) 545–559. doi:10.1017/jfm.2018.366.