ARTICLE IN PRESS
YACHA:1181

Appl. Comput. Harmon. Anal. • • • (• • • •) • • •–• • •

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Mixed Hölder matrix discovery via wavelet shrinkage and Calderón–Zygmund decompositions

Jerrod Ankenman [a], William Leeb [b,*]

[a] *Susquehanna International Group, Bala Cynwyd, PA 19004, United States*
[b] *PACM, Princeton University, Princeton, NJ 08544, United States*

A R T I C L E I N F O

Article history:
Received 13 January 2016
Received in revised form 6 December 2016
Accepted 23 January 2017
Available online xxxx
Communicated by Charles K. Chui

MSC:
62G08
42C40
26B35
05C05

Keywords:
Hölder
Mixed Hölder
Tensor product
Tree metric
Wavelet
Haar system
Wavelet shrinkage
Besov space
Calderón–Zygmund decomposition

A B S T R A C T

This paper concerns two related problems in the analysis of data matrices whose rows and columns are equipped with tree metrics. First is the problem of recovering a matrix that has been corrupted by additive noise. Under the assumption that the clean matrix exhibits a specific regularity condition, known as the mixed Hölder condition, we adapt the well-known Donoho–Johnstone wavelet shrinkage methods from classical nonparametric statistics to obtain estimators that are within a logarithmic factor of the minimax error rate with respect to mean squared error loss.

The second part of this paper develops a theory of Besov spaces on products of tree geometries. We show that matrices with small Besov norm can be written as a sum of a mixed Hölder matrix and a matrix with small support. Such decompositions are known as Calderón–Zygmund decompositions and are of general interest in harmonic analysis. The decompositions we establish impose fewer conditions on the function with small support than previous decompositions of this type while maintaining the same guarantees on the mixed Hölder matrix. As such, they are applicable to a greater variety of matrices and should find use in many data organization problems. As part of our analysis, we provide characterizations of the underlying Besov spaces using wavelets and other multiscale difference operators that are analogous to those from the classical Euclidean theory.

1. Introduction

This paper is concerned with matrix decompositions of the following form: if $f(x, y)$ is a matrix, by which we mean a function on the product of two discrete sets X and Y, we seek to write $f = g + b$, where g is a "good" matrix satisfying a certain regularity condition known as the mixed Hölder condition that we describe in Section 2, and b is a "bad" matrix that is nevertheless under control in some way.

* Corresponding author.
E-mail addresses: mathofpoker@gmail.com (J. Ankenman), wleeb@math.princeton.edu (W. Leeb).

http://dx.doi.org/10.1016/j.acha.2017.01.003
1063-5203/

Please cite this article in press as: J. Ankenman, W. Leeb, Mixed Hölder matrix discovery via wavelet shrinkage and Calderón–Zygmund decompositions, Appl. Comput. Harmon. Anal. (2017), http://dx.doi.org/10.1016/j.acha.2017.01.003

Such decompositions are encountered throughout analysis and its applications, such as in signal and image processing [1].

In Sections 1.1–1.6, we briefly introduce the high-level ideas used throughout this paper. In Section 1.7, we discuss the contributions of this paper.

1.1. Wavelets and multiresolution analysis

We give a brief summary of some relevant facts from wavelet theory. Of particular concern to us will be the notion of a *multiresolution analysis* of $L^2(\mathbb{R})$ [2–4]. One starts with a function $\phi(x)$, and considers all its dyadic dilates and integer translates, given by

$$\phi_{j,k}(x) = 2^{-j/2}\phi(2^{-j}x - k) \tag{1}$$

We define V_j as the linear span of the functions $\phi_{j,k}$ over all integers k. Under suitable conditions on ϕ, these spaces will be nested; that is, $V_j \subsetneq V_{j-1}$, or in other words, we can write $\phi(x)$ as a linear combination of the functions $\phi(2x - k)$; and their union is all of $L^2(\mathbb{R})$. In this case the system of subspaces V_j forms what is called a "multiresolution analysis" of $L^2(\mathbb{R})$, as each subspace captures activity at a certain dyadic scale, or resolution.

Wavelet analysis arises by looking at the orthogonal complement of V_j in V_{j-1}, which we denote by W_j. Given a multiresolution analysis as just described, one can construct a function $\psi(x)$ whose integer translates span W_0, and consequently where the functions $\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k)$ span W_j. The function ϕ is known as the "father wavelet", or "scaling function" and the function ψ is known as the "mother wavelet".

Perhaps the simplest example of such a system is the Haar system. Here, the father wavelet ϕ is the indicator function of the interval $[0, 1]$, and the mother wavelet is the function $\chi_{[0,1/2]} - \chi_{[1/2,1]}$. The space V_j is the span of indicator functions of dyadic intervals $[2^{-j}k, 2^{-j}(k+1)]$ for all integers k. It is very simple to generalize this particular multiresolution analysis to the setting of partition trees on abstract sets [5], as we will describe in more detail later.

1.2. The classical Besov spaces

Given a metric space (X, d), a natural way of measuring the variation of a function f defined on X is its Lipschitz norm, defined by

$$\sup_{x \neq y} \frac{f(x) - f(y)}{d(x, y)}. \tag{2}$$

If f is a differentiable function on $\mathbb{R}$, the Lipschitz norm (2) is equal to $\|f'\|_\infty$, the supremum of f's derivative. Expression (2), however, is defined for non-differentiable functions and makes sense in the abstract setting of any metric space.

A generalization of the Lipschitz norm is the Hölder norm, which replaces the metric $d(x, y)$ by $d(x, y)^\alpha$ for some parameter $\alpha > 0$. For functions on $\mathbb{R}$, this space is only non-trivial when $0 < \alpha \le 1$. The space of Hölder functions when α is strictly less than 1 has nicer algebraic properties than the Lipschitz space; in particular, the Hölder norm of a function can be characterized by the size of its wavelet coefficients. If we take a sufficiently nice wavelet basis $\{\psi_{j,k}\}$ of $\mathbb{R}^n$ (where $j \in \mathbb{Z}$ indexes the dyadic scale 2^{-j} and $k \in \mathbb{Z}$ the location), then the expression

$$\sup_{j,k} 2^{j(\alpha+1/2)} |\langle f, \psi_{j,k} \rangle| \tag{3}$$

Please cite this article in press as: J. Ankenman, W. Leeb, Mixed Hölder matrix discovery via wavelet shrinkage and Calderón–Zygmund decompositions, Appl. Comput. Harmon. Anal. (2017), http://dx.doi.org/10.1016/j.acha.2017.01.003