Adaptive Nonlinear Approximations

by
Geoffrey Davis

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS
COURANT INSTITUTE OF MATHEMATICAL SCIENCES
NEW YORK UNIVERSITY

SEPTEMBER 1994

Approved

Stéphane Mallat
Advisor

To my parents

Acknowledgements

I have been exceptionally fortunate in having Stéphane Mallat as my research advisor during my time at Courant. Not only have I learned a great deal of mathematics in my work with Stéphane, but just as importantly, I have learned much about being a mathematician. I thank him for his patient and generous guidance.

My colleagues Zhifeng Zhang, Francois Bergeaud, and Wen-Liang Hwang have made my research all the more enjoyable. I would like especially to thank Francois for his warm hospitality (and his *carte d'identité*!) during the fall I spent in Paris. I am also very grateful to Zhifeng for his Matching Pursuit Package software, which was used for numerical experiments involving the Gabor dictionary throughout the thesis. I would also like to thank Marco Avellaneda for many helpful discussions about stochastic processes.

Abstract

The problem of optimally approximating a function with a linear expansion over a redundant dictionary of waveforms is NP-hard. The greedy matching pursuit algorithm and its orthogonalized variant produce sub-optimal function expansions by iteratively choosing the dictionary waveforms which best match the function's structures. Matching pursuits provide a means of quickly computing compact, adaptive function approximations.

Numerical experiments show that the approximation errors from matching pursuits initially decrease rapidly, but the asymptotic decay rate of the errors is slow. We explain this behavior by showing that matching pursuits are chaotic, ergodic maps. The statistical properties of the approximation errors of a pursuit can be obtained from the invariant measure of the pursuit. We characterize these measures using group symmetries of dictionaries and using a stochastic differential equation model. These invariant measures define a noise with respect to a given dictionary. The dictionary elements selected during the initial iterations of a pursuit correspond to a function's coherent structures. The expansion of a function into its coherent structures provides a compact approximation with a suitable dictionary. We demonstrate a denoising algorithm based on coherent function expansions. We also introduce an algorithm for adapting a dictionary for efficiently decomposing a given class of functions.

Contents

List of Figures

Chapter 1

Introduction

1.1 Redundant Representations

The focus of this work is the problem of obtaining efficient representations of functions. Specifically, we seek to approximate functions with linear combinations of a small number of unit vectors from a family $\{g_\gamma\}_{\gamma\in\mathbf{\Gamma}}$ in a Hilbert space $\mathcal{H}$. For any $M > 0$, we want to minimize the error

$$\epsilon(M) = \|f - \sum_{\gamma\in I_M} \beta_\gamma g_\gamma\|$$

where $I_M \subset \mathbf{\Gamma}$ is an index set of cardinality M. Representations of this form are of central importance in numerous applications. Image compression requires efficient storage of functions $f(x,y)$ on the plane. If we can accurately approximate f with a linear combination of a small number of the vectors g_γ, then we need only store a small number of coefficients β_γ and indices γ. For numerical methods, such representations can reduce lengthy computations on f to a small number of computations performed on each g_γ in the expansion of f, enabling fast calculation. In pattern recognition applications, the g_γ in the expansion of f are interpreted as features of f. Compact expansions highlight the dominant features of f and allow f to be characterized by a few salient characteristics.

When $\{g_\gamma\}_{\gamma\in\mathbf{\Gamma}}$ is an orthonormal basis we can minimize the approximation error $\epsilon(M)$ by taking I_M to be the vectors corresponding to the largest M inner products $(|< f, g_\gamma >|)_{\gamma\in\mathbf{\Gamma}}$, since

$$\epsilon(M) = \sum_{\gamma\in\mathbf{\Gamma}-I_M} |< g_\gamma, f >|^2.$$

For the case that $\mathcal{H}$ is a space of finite dimension N and the set $\mathbf{\Gamma}$ contains a finite number P orthogonal vectors, the expansion which minimizes $\epsilon(M)$ is not difficult to compute and requires $O(PN)$ work.

Depending upon the basis and the space $\mathcal{H}$, it is possible to estimate the decay rate of the minimal approximation error $\epsilon_0(M) = \inf_{I_M} \epsilon(M)$ as M increases. For example, if $\{g_\gamma\}_{\gamma\in\mathbf{\Gamma}}$ is a wavelet basis, the rate of decay of $\epsilon_0(M)$ can be estimated for functions that belong to any particular Besov space. Conversely, the rate of decay of $\epsilon_0(M)$ characterizes the Besov space to which f belongs[15].

We can greatly improve these linear approximations to f by enlarging the collection $\{g_\gamma\}_{\gamma\in\mathbf{\Gamma}}$ beyond a basis. This enlarged, redundant family of vectors we call a dictionary. The advantage of redundancy in obtaining compact representations can be seen by considering the problem of representing a two-dimensional surface given by $f(x, y)$ on a subset of the plane, $I \times I$, where I is the interval $[0, 1]$. An adaptive square mesh representation of f in the Besov space $B^\alpha_q(L^q(I))$, where $\frac{1}{q} = \frac{\alpha+1}{2}$, can be obtained using a wavelet basis. This wavelet representation can be shown to be asymptotically near optimal in the sense that the decay rate of the error $\epsilon(M)$ is equal to the fastest decay attainable by a general class of non-linear transform-based approximation schemes [16][17]. Even these near-optimal representations are constrained by the fact that the decompositions are over a basis. The regular grid structure of the wavelet basis prevents the compact representation of many functions. For example, when f is equal to a basis wavelet at the largest scale, it can be represented exactly by a expansion consisting of a single element. However, if we translate this f slightly, then an accurate approximation can require many elements. One way to improve matters is to add to the set $\{g_\gamma\}_{\gamma\in\mathbf{\Gamma}}$, for example, the collection of all translates of the wavelets. The class of functions which can be compactly represented will then be translation invariant. We can do even better by expanding the dictionary to contain the extremely redundant set of all piecewise polynomial functions on arbitrary triangles.

Communicating in a natural language is another example of the use of compact representations in overcomplete sets. The English language is highly redundant as the heft of any thesaurus will show. Careful selection of words, however, allows precise, richly detailed information to be conveyed succinctly–consider a well-crafted *haiku*, for example.

Intriguing research in neurophysiology suggeststhat compact representations over a highly redundant set have a deep biological analog that is an integral component of human cognition. Information from the retina is thought to pass through a hierarchy of feature detectors in the cortex, each layer of which corresponds to a larger and more complex set of features from which fewer and fewer features are identified. The activity of cells in the retina corresponds roughly to the brightness and frequency of light at a particular location on the retina. At the next level, different ganglion cells are tuned to respond to particular local orientations of edges, local movement at a particular rate, and so on. It is postulated that the end result of this process is a characterization of the sensory information by as few active neurons as possible, perhaps as few as the 1,000 words which describe a picture[3] [4].

1.2 Practical Considerations

The first issue we must resolve is how to find compact expansions for a given function f. We require that finite linear combinations of dictionary vectors $\{g_\gamma\}_{\gamma\in\mathbf{\Gamma}}$ be dense in the space $\mathcal{H}$. Hence, it is always possible to obtain a linear expansion of any $f \in \mathcal{H}$. When there exist constants $A, B > 0$ such that for all $f \in \mathcal{H}$

$$A\|f\|^2 \leq \sum_{\gamma \in \mathbf{\Gamma}} |< f, g_\gamma >| \leq B\|f\|^2 \tag{1.1}$$

the collection $\{g_\gamma\}$ is called a *frame*. Frames were first introduced 40 years ago in [19] for use with nonharmonic Fourier series. They have received much attention in recent years as a tool for analyzing discrete wavelet transforms [13] [57]. The frame condition (1.1) implies that the linear operator $T : \mathcal{H} \rightarrow \mathcal{H}$ defined by

$$Tf = \sum_{\gamma \in \mathbf{\Gamma}} < f, g_\gamma > g_\gamma \tag{1.2}$$

is invertible. The inverse of T gives rise to a dual frame $\{\tilde{g}_\gamma\}_{\gamma \in \mathbf{\Gamma}} = \{T^{-1} g_\gamma\}_{\gamma \in \mathbf{\Gamma}}$, from which we obtain an explicit expansion formula for f,

$$f = \sum_{\gamma \in \mathbf{\Gamma}} < f, \tilde{g}_\gamma > g_\gamma. \tag{1.3}$$

For any other expansion $\sum \beta_\gamma g_\gamma = f$, the sum of the coefficients $\sum |\beta_\gamma|^2 > \sum |< f, \tilde{g}_\gamma >|^2$. These frame reconstructions utilize *all* g_γ's in the dictionary, however, and hence do not in general provide the compact representations we seek.

In chapter 2 we study the complexity of finding optimal M vector approximations to a function f, expansions of f which have minimal approximation error $\epsilon(M)$. We prove that in spaces of finite dimension N the problem of finding expansions in $\alpha_1 N \leq M \leq \alpha_2 N$ vectors from a redundant dictionary that minimize the error $\epsilon(M)$ is in general a fundamentally intractable problem–in fact, it is NP-hard.

Because of the difficulty of finding optimal approximations, two alternative expansion-finding strategies have emerged. The first is to use exhaustive search methods to find an exact solution to a sharply restricted and much easier problem, an approach taken by vector quantization and the best-basis algorithm. The second method is to use an iterative greedy algorithm to approximate optimal solutions of the general problem, a strategy employed by matching pursuit algorithms and their variants.

1.3 Constrained Expansions–Vector Quantization and Best-Basis

Vector quantization was introduced by Shannon in the 1940's as a device for obtaining information theoretical bounds [52]. In the last decade the advent of high speed computing has brought these techniques into wide use for data compression [26]. Shape-gain vector quantization [6] [50] is a type of quantization designed to approximate patterns in vectors which occur over a range of different gain values. In our framework, shape-gain vector quantization is equivalent to approximating a function f with the single term sum $\beta_\gamma g_\gamma$.

The vector g_γ is chosen from a large, highly redundant collection of unit vectors called a codebook. Because of the extremely small size of the expansions, quantization algorithms employ the brute force method of trying expansions using all vectors in the codebook to find the optimal one. The optimal expansion is given by $< f, g_\gamma > g_\gamma$ where g_γ is the vector which maximizes $|< f, g_\gamma >|$; the magnitude of the relative error is

$$\frac{\epsilon(1)}{\|f\|^2} = 1 - \frac{|< f, g_\gamma >|^2}{\|f\|^2}. \tag{1.4}$$

To ensure a small error the g_γ's must form a very dense set on the surface of the unit sphere. The small number of terms in the expansions therefore places a sharp limit on the dimension of the space from which functions can be approximated with any degree of accuracy, since the size of the codebooks needed to cover the sphere with a given density increases exponentially with the dimension of the space. To expand large dimensional signals, such as digital audio recordings or images, the signals are first segmented into low-dimensional components, and these components are then quantized. This segmentation of the signal is equivalent to a restricting the dictionary to a collection of mutually orthogonal blocks of vectors. With such a dictionary, the expansions can only represent efficiently those structures that are limited to a single low-dimensional partition. Structures that extend across the partitions require many dictionary vectors for accurate representation.

The best-basis algorithm of Coifman and Wickerhauser [9] performs function expansions over orthogonal bases from a carefully constructed dictionary. For an N dimensional space, the best basis dictionary contains $N \log_2 N$ functions of the form

$$g_\gamma(t) = 2^{\frac{j}{2}} w_n(2^j t - k), \quad n \in \mathbf{N},\ j, k \in \mathbf{Z} \tag{1.5}$$

called wavelet packets. The parameters k and j are translation and scaling parameters, and n corresponds roughly to a modulation. This set of wavelet packets contains over 2^N different orthonormal bases of $\mathcal{H}$, including orthonormal wavelet bases and an orthogonalized analog of the window Fourier transform. The structure of the dictionary can be utilized to allow decompositions of functions to be computed in $O(N \log N)$ time. This fast performance is a result of the particular structure of the wavelet packet dictionary, however, and the algorithm does not generalize to other dictionaries. Hence, the type of expansions that can be performed is restricted. Another limitation of the algorithm is that the expansions are constrained to orthonormal bases. This restriction can preclude more efficient expansions. The presence of strong transients within a signal, for example, can mask the presence of nearby portions of the signal with different time-frequency behaviors by causing the algorithm to choose a local basis that is well-suited to decomposing only the transients.

1.4 General Expansions with Greedy Algorithms

A class of iterative algorithms for performing expansions over redundant dictionaries, called matching pursuits, have been independently developed in signal processing [58], statistics

[32][20][30], and control theory applications [7][5]. An excellent overview is provided in [58]. Matching pursuits are greedy algorithms; rather than finding a globally optimal expansion (which we show is an NP-hard problem), at each step of operation they find an optimal one-element expansion. The residual of this one element expansion is then expanded in the next iteration, and so on. We review a fast, parallelizable version of matching pursuit algorithm due to [58] in chapter 3.

Because the pursuit is a multi-stage process, the dictionaries used for the expansions do not need to be as enormous as those used for single stage vector quantization. Matching pursuits are therefore capable of decomposing functions in very high dimensional spaces without requiring enormous computational resources or partitioning into orthogonal subspaces. Moreover, because high dimensional spaces do not need to be partitioned, the dictionaries can include structures which are much more delocalized than can single stage vector quantization codebooks.

Unlike the best-basis algorithm, matching pursuits can use arbitrary dictionaries. In section 3.3 we describe an application of matching pursuits for performing adaptive time-frequency decompositions of signals. For this application we use as a dictionary a family of vectors which are optimally localized in the time-frequency plane, a collection of translated, modulated, and scaled Gaussians. Wavelet packets have poor frequency localization [42] and are thus much less efficient for performing this task. The constraint that the best-basis decomposition be an orthogonal basis is an additional hindrance to performance when signals are non-stationary. The basis constraint imposes a structure on the decomposition which can prevent adaptation to local signal structures when strong features of very different time-frequency localization are nearby. The greedy expansions of matching pursuits, in contrast, are locally adapted to the time-frequency localization of signal structures.

The question now arises of what we sacrifice in using these non-optimal greedy expansions. For dictionaries consisting of an orthonormal basis, the matching pursuit expansion for a function f is precisely the optimal expansion obtained above by using the M dictionary elements with the largest inner products $|< f, g_\gamma >|$. For more general dictionaries the matching pursuit expansions are not optimal. In fact, when no two elements of a dictionary for a finite dimensional space are orthogonal, matching pursuit expansions are not only non-optimal, they do not converge in a finite number of iterations except on a set of measure 0. We introduce an orthogonal matching pursuit which converges in finite steps in finite dimensional spaces, and we compare the performance and complexity of the non-orthogonal and orthogonal pursuit.

We examine the asymptotic behavior of matching pursuits in order to better understand their convergence properties. We prove that matching pursuits possess chaotic properties and that for a class of dictionaries with a group invariance the asymptotic approximation errors can be viewed as the realizations of a stationary white process called dictionary noise. The asymptotic convergence of the pursuit can be quite slow, but our numerical experiments show that given a suitable dictionary, function expansions initially converge very quickly. Thus, when the number of terms in the expansion is not too large, matching pursuits provide efficient approximations.

1.5 Dictionaries

An additional issue we must resolve is how to determine an appropriate dictionary over which to perform the expansions. Even an optimal expansion will not provide compact function representations if an unsuitable dictionary is used. For example, a discrete Fourier basis is a poor choice for expanding functions containing discontinuities. Similarly, we would not want to use a Haar wavelet basis for approximating smooth functions.

The problem of finding an appropriate dictionary is essentially one of finding a set of canonical features which characterize the functions we wish to decompose. Experiments show that animals reared in environments lacking certain types of visual stimuli have few if any neurons that respond to the missing features when they are presented later in life. This shows that the "dictionary" of features used by the cortex to encode visual sensory data is at least partially learned. We develop an algorithm which, like the cortex, iteratively adapts a dictionary to provide efficient representations for a training set of data.[3]

We examine the specific problem of optimizing a dictionary for approximating the realizations of a given random process. Many types of physical data can be viewed as such realizations. An important class of dictionaries we study are those which possess a group structure, such as translation or modulation invariance. Such dictionaries are *a priori* well-suited to decomposing realizations of a process which possesses similar invariances. Apart from establishing such macroscopic properties as translation invariance, the problem of finding an optimal dictionary is a difficult one. Dictionary optimization has been well-studied in the context of vector quantization, i.e. in the case of expansions consisting of a single vector. The generalized Lloyd algorithm [34] is a standard method for optimizing dictionaries for vector quantization. We present a modified version of the Lloyd algorithm to iteratively optimize a dictionary for approximating the realizations of a particular random process.

1.6 Outline of Thesis

In this thesis we address three central issues for performing function expansions over redundant dictionaries.

1. How can we efficiently obtain expansions over redundant dictionaries which minimize the approximation error $\epsilon(M)$, and what is the computational complexity of obtaining such expansions?

2. For what types of functions can we obtain compact representations with a given dictionary, and how can we characterize the approximation errors from such a scheme?

3. How can we find a dictionary which is optimal for a given class of functions?

In chapter 2 we prove that the problem of finding optimal function expansions over a redundant dictionary is NP-hard. We show that the minimal approximation error criterion leads to intrinsically unstable expansions which accounts for some of the difficulty in finding these expansions.

As a result of these complexity results, we turn to sub-optimal methods for finding expansions. In chapter 3 we review the matching pursuit algorithm of [39], and in chapter 4.1 we introduce an orthogonalized version of the algorithm which has improved convergence properties. We compare the complexity and performance of these two algorithms for a set of speech data with a dictionary of time-frequency atoms.

In chapters 6 and 7 we study the asymptotic behavior of the residuals from a matching pursuit. In our numerical experiments we find that the rate of decay of the approximation error $\epsilon(M)$ decreases as M becomes large. These observations are explained by showing that a matching pursuit is a chaotic map which has an ergodic invariant measure. The proof of chaos is given for a particular dictionary in a low-dimensional space, and we show numerical results which indicate that higher dimensional matching pursuits are also ergodic maps. When M is small, the matching pursuit provides an efficient function expansion into what we call "coherent" structures. The error incurred by truncating function expansions when the convergence rate $\epsilon(M)$ becomes small corresponds to the realization of a process which is characterized by the invariant measure called "dictionary noise." The properties of invariant measures are studied for the particular case of dictionaries that are invariant under the action of a group of operators, and a stochastic model of the evolution of the residues is developed for a dictionary which is composed of a discrete Dirac basis and discrete Fourier basis.

Finally, in chapter 8 we address the problem of adapting a dictionary for decomposing realizations of a particular random process. We use a modified version of the Lloyd algorithm of vector quantization to develop an algorithm for iteratively optimizing a dictionary for matching pursuit and orthogonal matching pursuit expansions.

Chapter 2

Complexity of Optimal Approximation

Let $\mathcal{H}$ be a Hilbert space. A dictionary for $\mathcal{H}$ is a family $\mathcal{D} = \{g_\gamma\}_{\gamma \in \Gamma}$ of unit vectors in $\mathcal{H}$ such that finite linear combinations of the g_γ are dense in $\mathcal{D}$. The smallest possible dictionary is a basis of $\mathcal{H}$; general dictionaries are redundant families of vectors. Vectors in $\mathcal{H}$ do not have unique representations as linear sums of redundant dictionary elements. We prove below that for a redundant dictionary we must pay a high computational price to find an expansion with M dictionary vectors that yields the minimum approximation error.

Definition 2.1 *Let $\mathcal{D}$ be a dictionary of functions in an N-dimensional Hilbert space $\mathcal{H}$. Let $\epsilon > 0$ and $M \in \mathbf{N}$. For a given $f \in \mathbf{R}^N$ an (ϵ, M)*-**approximation** *is an expansion*

$$\tilde{f} = \sum_{i=1}^{M} \beta_i g_{\gamma_i}, \tag{2.1}$$

where $\beta_i \in \mathbf{C}$ and $g_{\gamma_i} \in \mathcal{D}$, for which

$$||\tilde{f} - f|| < \epsilon.$$

An **M-optimal approximation** *is an expansion that minimizes $||\tilde{f} - f||$.*

If our dictionary consists of an orthogonal basis, we can obtain an M-optimal approximation for any $f \in \mathcal{H}$ by computing the inner products $\{< f, g_\gamma >\}_{\gamma \in \mathbf{\Gamma}}$, and sorting the dictionary elements so that $|< f, g_{\gamma_i} >| \geq |< f, g_{\gamma_{i+1}} >|$. The signal $\tilde{f} = \sum_{i=1}^{M} < f, g_{\gamma_i} > g_{\gamma_i}$ is then an M-optimal approximation to f. In an N dimensional space, computing the inner products requires $O(N^2)$ operations and sorting $O(N \log N)$ so the overall algorithm is $O(N^2)$.

For general redundant dictionaries, the following theorem proves that finding M-optimal approximations is computationally intractible.

Theorem 2.1 *Let $\mathcal{H}$ be an N dimensional Hilbert space. Let $k \geq 1$ and let $\mathcal{D}_N$ be the set of all dictionaries for $\mathcal{H}$ that contain $O(N^k)$ vectors. Let $0 < \alpha_1 < \alpha_2 < 1$ and $M \in \mathbf{N}$ such that $\alpha_1 N \leq M \leq \alpha_2 N$. The* **($\epsilon$, M)-approximation problem***, determining for any given $\epsilon > 0, \mathcal{D} \in \mathcal{D}_N$, and $f \in \mathcal{H}$, whether an (ϵ, M)-approximation exists, is NP-complete. The* **M-optimal approximation problem***, finding the optimal M-approximation, is NP-hard.*

The theorem does not imply that the M-approximation problem is intractible for *specific* dictionaries $\mathcal{D} \in \mathcal{D}_N$. Indeed, we saw above that for orthonormal dictionaries, the problem can be solved in polynomial time. Rather, we mean that if we have an algorithm which finds the optimal approximation to any given $f \in \mathbf{R}^N$ for *any* dictionary $\mathcal{D} \in \mathcal{D}_N$, the algorithm decides an NP-hard problem.

Note: in our computations we restrict f, the elements of the dictionaries, and their coefficients to floating point representations of $\Theta(N^m)$ bits for some fixed m [53]. This restriction does not substantially affect the proof of NP-completeness, because the problems that must be solved for the proof are discrete and unaffected by small perturbations.

Proof: For any ϵ we can solve the (ϵ, M)-approximation problem by first solving the M-optimal approximation problem, computing $\epsilon_{min} = \|\tilde{f} - f\|$, and then checking whether $\epsilon_{min} < \epsilon$. Hence the M-optimal approximation problem must be at least as hard as the (ϵ, M)-approximation problem. Proving that the (ϵ, M)-approximation problem is NP-complete thus implies that the M-optimal approximation problem is NP-hard. The (ϵ, M)-approximation problem is in NP, because we can verify in polynomial time that $\|\tilde{f} - f\| < \epsilon$ once we are given the set of M elements and their coefficients. To prove that it is NP-complete we prove that it is as hard as the exact cover by 3-sets problem, a problem which is known to be NP-complete.

Definition 2.2 *Let X be a set containing $N = 3M$ elements, and let $\mathcal{C}$ be a collection of 3-element subsets of X. The* **exact cover by 3-sets** *problem is to decide whether $\mathcal{C}$ contains an exact cover for X, i.e. to determine whether $\mathcal{C}$ contain a subcollection $\mathcal{C}'$ such that every member of X occurs in exactly one member of $\mathcal{C}'$? [23]*

Lemma 2.1 *We can transform in polynomial time any instance $(X, \mathcal{C})$ of the exact cover by 3-sets problem of size $|X| = 3M$ into an equivalent instance of the (ϵ, M)-approximation problem with a dictionary of size $O(N^3)$ in an N-dimensional Hilbert space.*

This lemma implies that if we can solve the (ϵ, M)-approximation problem for $M = N/3$, we can also solve an NP-complete problem so the approximation problem must be NP-complete as well. It thus gives a proof of the theorem for $M = \frac{N}{3}$.

Proof: Let $\mathcal{H}$ be an N dimensional space with an orthonormal basis $\{e_i\}_{1 \leq i \leq N}$. For notational convenience we suppose that X is the set of $N = 3M$ integers between 1 and N. Let $\mathcal{C}$ be a collection of 3-element subsets of X. To any subset of K integers $S \subseteq X$

we associate a unit vector in $\mathcal{H}$ defined by

$$T(S) = \sum_{i \in S} \frac{e_i}{\sqrt{K}}. \tag{2.2}$$

Let $\mathcal{D}$ be the dictionary of $\mathcal{H}$ defined by

$$\mathcal{D} = \{T(S_i) : S_i \in \mathcal{C}\}, \tag{2.3}$$

where the S_i's are the three-element subsets of X contained in $\mathcal{C}$. Since $\mathcal{C}$ contains at most $\left(\begin{array}{c} N \\ 3 \end{array} \right) = O(N^3)$ three-element subsets of X, this transformation can be done in polynomial time.

We now show that solving the (ϵ, M)-approximation problem for

$$f = T(X) = \sum_{i=1}^{N} \frac{e_i}{\sqrt{N}} \tag{2.4}$$

and $\epsilon < \frac{1}{\sqrt{N}}$ is equivalent to solving the exact cover by 3-Sets problem $(X, \mathcal{C})$. Suppose $\mathcal{C}$ contains an exact cover $\mathcal{C}'$ for X. Then

$$\| \sum_{S_i \in \mathcal{C}'} \sqrt{\frac{3}{N}} T(S_i) - f \| = 0. \tag{2.5}$$

Since there are $M = \frac{1}{3}N$ such S_i's, the approximation problem has a solution. Thus, a solution to the exact cover problem implies a solution to the approximation problem.

Conversely, suppose the (ϵ, M)-approximation problem has a solution for $\epsilon < \frac{1}{\sqrt{N}}$. There exist M three-element sets $S_i \in \mathcal{C}$ and M coefficients β_n such that

$$\| \sum_{n=1}^{M} \beta_n T(S_n) - f \| < \frac{1}{\sqrt{N}}.$$

The inner product of each basis vector $\{e_i\}_{1 \le i \le N}$ with $\sum_{n=1}^{M} \beta_n T(S_n)$ must be non-zero, for otherwise we would have $\| \sum_{i=1}^{n} \beta_i T(S_i) - f \| \ge \frac{1}{\sqrt{N}}$ (recall that all components of f are equal to $\frac{1}{\sqrt{N}}$). Since each $T(S_i)$ has non-zero inner products with exactly three bases vectors and $N = 3M$, the M sets $(S_i)_{1 \le i \le M}$ do not intersect and thus define an exact 3-set cover of X. This proves that a solution to the approximation problem implies a solution to the Exact Cover problem, which finishes the proof of the lemma.

□

We have proved that the (ϵ, M)-approximation problem is NP-complete for $\alpha_1 = \alpha_2 = \frac{1}{3}$ and dictionaries of size $O(N^3)$. We now extend the result to arbitrary $0 < \alpha_1 < \alpha_2 < 1$ and dictionaries of size $O(N^k)$ for $k > 1$. Let $(X, \mathcal{C})$ be an instance of the exact cover by 3-sets problem where X is a set of n elements. Following lemma 2.1, we construct an equivalent (ϵ, M)-approximation problem on a $3n$-dimensional space $\mathcal{H}_1$. We then embed this approximation problem in a larger Hilbert space $\mathcal{H}$ in order to satisfy the dictionary size and expansion length constraints. In $\mathcal{H}_2$, the orthogonal complement of $\mathcal{H}_1$ in $\mathcal{H}$, we construct a $(0, \alpha_2 N - n)$-approximation problem which has a unique solution. The combined approximation problem in $\mathcal{H}$ will be equivalent to the exact cover problem and will have the requisite M and dictionary size.

Let N be the smallest integer such that $N \geq 3n$, $N^k \geq |\mathcal{C}|$, and $\alpha_2 N \geq 3n$. This N is bounded by a polynomial in n, since $|\mathcal{C}| \leq n^3$. Let $\mathcal{H}$ be an N-dimensional Hilbert space and let $\{e_i\}_{1 \leq i \leq N}$ be an orthonormal basis of $\mathcal{H}$. Let $\mathcal{H}_1$ be the subspace of $\mathcal{H}$ spanned by the vectors $\{e_i\}_{1 \leq i \leq 3n}$ (recall that $N \geq 3n$). We map subsets of X to $\mathcal{H}_1$ using (2.2) as we did in lemma 2.1 and we define

$$f_1 = \sqrt{3}\beta T(X) = \beta \sum_{i=1}^{3n} e_i \tag{2.6}$$

where β is a constant we will define below, and we set

$$\mathcal{D}_1 = \{T(S_i) : s_i \in \mathbf{C}\}. \tag{2.7}$$

This mapping can be done in time bounded by a polynomial in n because $|\mathcal{C}| \leq n^3$. From the proof of lemma 2.1 we see that the function f_1 can be approximated with an error of less than β using n vectors from $\mathcal{D}_1$ if and only if X has an exact cover, and it cannot be approximated to within β using less than n vectors.

We now create a second approximation problem in $\mathcal{H}_2$, the orthogonal complement of $\mathcal{H}_1$ in $\mathcal{H}$, so that we can control the size of the expansion. We define

$$f_2 = \beta \sum_{i=3n+1}^{\lfloor \alpha_2 N \rfloor + 2n} e_i \tag{2.8}$$

and

$$\mathcal{D}_2 = \{e_i : 3n < i \leq \lfloor \alpha_2 N \rfloor + 2n\} \tag{2.9}$$

This construction, too, can be done in polynomial time in n, since N is bounded by a polynomial in n. Approximating f_2 to within β is only possible if entire set of $\lfloor \alpha_2 N \rfloor - n$ vectors $\mathcal{D}_2$ is contained in the expansion.

The approximation problem equivalent to $(X, \mathcal{C})$ is formed by setting

$$\begin{aligned} f &= f_1 + f_2 \\ \mathcal{D} &= \mathcal{D}_1 \cup \mathcal{D}_2 \end{aligned} \tag{2.10}$$

and choosing a positive $\epsilon < \beta$. We take $\beta = \frac{1}{\sqrt{\lfloor \alpha_2 N \rfloor + 2n}}$ so f is a unit vector. The combined dictionary contains $|\mathcal{C}| + \lfloor \alpha_2 N \rfloor - n < 2N^k$ vectors, so it is of $O(N^k)$ by our choice of N.

To prove the equivalence, we first suppose that $\mathcal{C}$ contains an exact cover $\mathcal{C}'$ for X. Then

$$||\sqrt{3}\beta \sum_{S_i \in \mathcal{C}'} T(S_i) + \beta \sum_{i=3n+1}^{\lfloor \alpha_2 N \rfloor + 2n} e_i - f|| = 0. \tag{2.11}$$

The first sum contains n terms and the second contains $\lfloor \alpha_2 N \rfloor - n$ terms, so the total number of terms M in the expansion is $\lfloor \alpha_2 N \rfloor$ which lies between $\alpha_1 N$ and $\alpha_2 N$ for N sufficiently large. Hence a solution to the exact cover problem implies a solution to the (ϵ, M)-approximation problem.

Suppose that $\mathcal{C}$ does not contain an cover for X. We can partition the error from our constructed approximation problem,

$$E = f - \sum_{i=1}^{M} \beta_i g_{\gamma_i} \tag{2.12}$$

into a sum of its projection E_1 onto $\mathcal{H}_1$ and its projection E_2 onto $\mathcal{H}_2$. The two subspaces are orthogonal, so $||E||^2 = ||E_1||^2 + ||E_2||^2$. We must therefore have $||E_1||^2 < \beta$ and $||E_2|| < \beta$. Now to obtain $||E_2|| < \beta$ we must include all $\lfloor \alpha_2 N \rfloor - n$ vectors from $\mathcal{D}_2$ in the expansion. We can therefore have at most n vectors from $\mathcal{D}_1$ in the expansion. From above, we have that an n-vector approximation to f_1 with error less than β is not possible without an exact cover. Hence, no solution to the (ϵ, M)-approximation problem exists. This proves that no solution to the exact cover problem implies no solution to the (ϵ, M)-approximation problem, thus proving the theorem.

□

A corollary of theorem 2.1 shows that finding approximations which have a minimum length for a given error tolerance is also intractable.

Corollary 2.1 *Let $\mathcal{H}$ be an N dimensional Hilbert space. Let $k \geq 1$ and let $\mathcal{D}_N$ be the set of all dictionaries for $\mathcal{H}$ that contain $O(N^k)$ vectors. Let $\epsilon > 0$. The* **ϵ-shortest approximation problem** *is to find the smallest M such that a linear combination of M dictionary elements*

$$\tilde{f}_\epsilon = \sum_{i=1}^{M} \beta_i g_{\gamma_i},$$

satisfies $||\tilde{f}_\epsilon - f|| < \epsilon(N)$. The ϵ-shortest approximation problem is NP-hard.

Proof: We prove that the problem is NP-hard by showing that we can use solutions of the problem to solve an NP-complete problem in polynomial time. Suppose we wish to decide the (ϵ, M)-approximation problem for some $\epsilon, \alpha_1, \alpha_2, f, D$, and N. We first solve the ϵ-shortest approximation problem to find the smallest number of dictionary elements M required to approximate f to within ϵ. If $M \leq \alpha_2 N$ then the (ϵ, M)-approximation problem has a solution. If not, there is no such (ϵ, M)-approximation.

□

Remarks: A problem related to the exact cover by 3-sets problem is the minimum cost cover problem. For X and $\mathcal{C}$ as in the 3-set problem, a cover of X is any collection of sets from $\mathcal{C}$ such that each element of X appears in at least one of the sets. The cost of a cover is the number of sets from $\mathcal{C}$ which make up the cover. The *minimum cost cover problem* is to find the cover for X with the smallest cost. When an exact cover for X exists, it is the minimum cost cover, so finding the minimum cost cover is at least as hard as the exact cover by 3-sets problem. The problem is NP-hard but is not in NP, because we cannot verify in polynomial time whether a given cover is minimal. A standard means of finding low-cost covers for X is to use a greedy method , and the ratio of the size of the greedily obtained cover to the size of the minimum cover can be shown to be bounded [12] [31] [35]. In the next section, we will describe a greedy method for approximating solutions to the M-optimal approximation problem

The optimal approximation criterion of definition 2.1 has number of undesirable properties which are partly responsible for its NP-completeness. The elements contained in the expansions are unstable in that functions which are only slightly different can have optimal expansions containing completely different dictionary elements. The expansions also lack an optimal substructure property. The expansion in M elements with minimal error does not necessarily contain an expansion in $M-1$ elements with minimal error. The expansions can therefore not be progressively refined. Finally, depending upon the dictionary, the coefficients of optimal approximations can exhibit instability in that the expansion coefficients β_i of the M-optimal approximation (2.1) to a vector f can have

$$\sum_{i=1}^{M} |\beta_i|^2 >> ||f||^2.$$

Consider the case when $\mathcal{H} = \mathbf{R}^3$, $f = (1,1,1)$, and $\mathcal{D} = \{e_1, e_2, e_3, v\}$, where the e_i's are the Euclidean basis of $\mathbf{R}^3$ and $v = \frac{e_1+\epsilon f}{||e_1+\epsilon f||}$. The M-optimal approximation to f for $M = 2$ is

$$\tilde{f} = \frac{||e_1 + \epsilon f||}{\epsilon} v - \frac{1}{\epsilon} e_1, \tag{2.13}$$

so we see that $\sum_{i=1}^{M} |\beta_i|^2$ can be made arbitrarily large. In the next section we describe an approximation algorithm based on a greedy refinement of the vector approximation, that maintains an energy conservation relation which guarantees stability.

Chapter 3

Matching Pursuits

A matching pursuit is a greedy algorithm that progressively refines the signal approximation with an iterative procedure instead of solving the optimal approximation problem. In section 3.1 we review this adaptive approximation procedure due to [58]. Section 3.2 describes a fast numerical implementation, and section 3.3 describes an application to an adaptive time-frequency decomposition. In the next chapter we introduce an orthogonalized version of the pursuit and compare the performance and complexity of the two algorithms.

3.1 Non-Orthogonal Matching Pursuits

Let $\mathcal{D} = \{g_\gamma\}_{\gamma \in \mathbf{\Gamma}}$ be a dictionary of vectors with unit norm in a Hilbert space $\mathcal{H}$. Let $f \in \mathcal{H}$. The first step of a matching pursuit is to approximate f by projecting it on a vector $g_{\gamma_0} \in \mathcal{D}$

$$f = < f, g_{\gamma_0} > g_{\gamma_0} + Rf. \tag{3.1}$$

Since the residue Rf is orthogonal to g_{γ_0},

$$\|f\|^2 = |< f, g_{\gamma_0} >|^2 + \|Rf\|^2. \tag{3.2}$$

We minimize the norm of the residue by choosing g_{γ_0} which maximizes $|< f, g_\gamma >|$. In infinite dimensions, the supremum of $|< f, g_\gamma >|$ may not be attained, so we choose g_{γ_0} such that

$$|< f, g_{\gamma_0} >| \geq \alpha \sup_{\gamma \in \Gamma} |< f, g_\gamma >|, \tag{3.3}$$

where $\alpha \in (0, 1]$ is an optimality factor. The vector g_{γ_0} is chosen from the set of dictionary vectors that satisfy (3.3), with a choice function whose properties vary depending upon the application.

The pursuit iterates this procedure by subdecomposing the residue. Let $R^0 f = f$. Suppose that we have already computed the residue $R^k f$. We choose $g_{\gamma_k} \in \mathcal{D}$ such that

$$|< R^k f, g_{\gamma_k} >| \geq \alpha \sup_{\gamma \in \mathbf{\Gamma}} |< R^k f, g_\gamma >| \tag{3.4}$$

and project $R^k f$ on g_{γ_k}

$$R^{k+1} f = R^k f - < R^k f, g_{\gamma_k} > g_{\gamma_k}. \tag{3.5}$$

The orthogonality of $R^{k+1} f$ and g_{γ_k} implies

$$||R^{k+1} f||^2 = ||R^k f||^2 - | < R^k f, g_{\gamma_k} > |^2. \tag{3.6}$$

By summing (3.5) for k between 0 and $n-1$ we obtain

$$f = \sum_{k=0}^{n-1} < R^k f, g_{\gamma_k} > g_{\gamma_k} + R^n f. \tag{3.7}$$

Similary, summing (3.5) for k between 0 and $n-1$ yields

$$||f||^2 = \sum_{k=0}^{n-1} |< R^k f, g_{\gamma_k} >|^2 + ||R^n f||^2. \tag{3.8}$$

The residue $R^n f$ is the approximation error of f after choosing n vectors in the dictionary and the energy of this error is given by (3.8). For any $f \in \mathcal{H}$, the convergence of the error to zero is shown [58] to be a consequence of a theorem proved by Jones [36]

$$\lim_{n \to \infty} ||R^n f|| = 0. \tag{3.9}$$

Hence

$$f = \sum_{k=0}^{\infty} < R^k f, g_{\gamma_k} > g_{\gamma_k}, \tag{3.10}$$

$$||f||^2 = \sum_{k=0}^{\infty} |< R^k f, g_{\gamma_k} >|^2. \tag{3.11}$$

In infinite dimensions, the convergence rate of this error can be extremely slow. In finite dimensions, let us prove that the convergence is exponential. For any vector $e \in \mathcal{H}$, we define

$$\lambda(e) = \sup_{\gamma \in \mathbf{\Gamma}} |< \frac{e}{||e||}, g_\gamma >|.$$

For simplicity, for the remainder of the thesis we will take the optimality factor α to be 1 for finite dimensional spaces unless otherwise specified. Hence, the chosen vector g_{γ_k} satisfies

$$\lambda(R^k f) = \frac{|< R^k f, g_{\gamma_k} >|}{||R^k f||}.$$

Equation (3.6) thus implies that

$$\|R^{k+1}f\|^2 = \|R^k f\|^2(1 - \lambda^2(R^k f)). \tag{3.12}$$

Hence norm of the residue decays exponentially with a rate equal to $-\frac{1}{2}\log(1 - \lambda^2(R^k f))$. Since $\mathcal{D}$ contains at least a basis of $\mathcal{H}$ and the unit sphere of $\mathcal{H}$ is compact in finite dimensions, we can derive [58] that there exists $\lambda_{min} > 0$ such that for any $e \in \mathcal{H}$

$$\lambda(e) \geq \lambda_{min}. \tag{3.13}$$

Equation (3.12) thus proves that the energy of the residue $R^k f$ decreases exponentially with a minimum decay rate equal to $-\frac{1}{2}\log(1 - \lambda^2_{min})$.

These non-orthogonal matching pursuits are similar in spirit to a class of iterative algorithms for estimating conditional expectations called projection pursuits proposed by [32] [54] and [55] and implemented by [21].

The central problem is to estimate the conditional expectation of a real-valued random variable Y with respect to an $\mathbf{R}^N$-valued random variable X. Specifically, for any $\mathbf{x} \in \mathbf{R}^N$ we would like to compute the expectation

$$f(x) = E(Y|X = \mathbf{x}) \tag{3.14}$$

from K observations of the variables $(X_1, Y_1), \ldots (X_K, Y_K)$. The algorithm works by iteratively projecting the function f onto a series of ridge functions, to obtain an expansion of the form

$$f(x) = \sum_{j=1}^{\infty} g_j(\mathbf{a}_j^T \mathbf{x}). \tag{3.15}$$

This projection pursuit algorithm was proved to converge strongly in [36], and from this result the proof of the convergence of the non-orthogonal matching pursuit is derived. The projection pursuits differ significantly from matching pursuits in that the function $f(x)$ is not known exactly, so its applicability and the numerical considerations for its implementation are quite different.

3.2 Numerical Implementation of Matching Pursuits

We suppose that $\mathcal{H}$ is a finite dimensional space and $\mathcal{D}$ a dictionary with a finite number of vectors. The optimality factor α is set to 1.

The matching pursuit is initialized by computing the inner products $\{< f, g_\gamma >\}_{\gamma \in \mathbf{\Gamma}_\alpha}$ and we store these inner products in an open hash table [11], where they are partially sorted. The algorithm is defined by induction as follows. Suppose that we have already computed $\{< R^n f, g_\gamma >\}_{\gamma \in \mathbf{\Gamma}}$, for $n \geq 0$. We must first find g_{γ_n} such that

$$|< R^n f, g_{\gamma_n} >| = \sup_{\gamma \in \mathbf{\Gamma}} |< R^n f, g_\gamma >|. \tag{3.16}$$

Since all inner products are stored in an open hash table, this requires $O(1)$ operations on average. Once g_{γ_n} is selected, we compute the inner product of the new residue $R^{n+1}f$ with all $g_\gamma \in \mathcal{D}$ using an updating formula derived from equation (3.5)

$$< R^{n+1}f, g_\gamma >=< R^n f, g_\gamma > - < R^n f, g_{\gamma_n} > < g_{\gamma_n}, g_\gamma > . \tag{3.17}$$

Since we have already computed $< R^n f, g_\gamma >$ and $< R^n f, g_{\gamma_n} >$, this update requires only that we compute $< g_{\gamma_n}, g_\gamma >$. Dictionaries are generally built so that few such inner products are non-zero, and non-zero inner products are either precomputed and stored or computed with a small number of operations. Suppose that the inner product of any two dictionary elements can be obtained with $O(I)$ operations and that there are $O(Z)$ non-zero inner products. Computing the products $\{< R^{n+1}f, g_\gamma >\}_{\gamma \in \mathbf{\Gamma}}$ and storing them in the hash table thus requires $O(IZ)$ operations. The total complexity of P matching pursuit iterations is thus $O(PIZ)$.

3.3 Application to Dictionaries of Time-Frequency Atoms

Signals such as sound recordings contain structures that are well localized both in time and frequency. This localization varies depending upon the sound, which makes it difficult to find a basis that is *a priori* well adapted to all components of the sound recording. Dictionaries of time-frequency atoms include waveforms with a wide range of time-frequency localization are thus much larger than a single basis. Such dictionaries are generated by translating, modulating, and scaling a single real window function $g(t) \in L^2(\mathbf{R})$. We suppose that that $g(t)$ is even, $||g|| = 1$, $\int g(t)dt \neq 0$, and $g(0) \neq 0$. We denote $\gamma = (s, u, \xi)$ and

$$g_\gamma(t) = \frac{1}{\sqrt{s}} g(\frac{t-u}{s}) e^{i\xi t}. \tag{3.18}$$

The time-frequency atom $g_\gamma(t)$ is centered at $t = u$ with a support proportinal to s. Its Fourier transform is

$$\hat{g}_\gamma(\omega) = \sqrt{s}\hat{g}(s(\omega - \xi))e^{-i(\omega-\xi)u}, \tag{3.19}$$

and the Fourier transform is centered at $\omega = \xi$ and concentrated over a domain proportional to $\frac{1}{s}$. For small values of s the atoms are well localized in time but poorly localized in frequency; for large values of s vice versa.

The dictionary of time-frequency atoms $\mathcal{D} = (g_\gamma(t))_{\gamma \in \Gamma}$ is a very redundant set of functions that includes both window Fourier frames and wavelet frames [13]. When the window function g is the Gaussian $g(t) = 2^{1/4}e^{-\pi t^2}$, the resulting time-frequency atoms are Gabor functions, and have optimal localization in time and in frequency. A matching pursuit decomposes any $f \in \mathbf{L}^2(\mathbf{R})$ into

$$f = \sum_{n=0}^{+\infty} < R^n f, g_{\gamma_n} > g_{\gamma_n}, \tag{3.20}$$

where the scales, position and frequency $\gamma_n = (s_n, u_n, \xi_n)$ of each atom

$$g_{\gamma_n}(t) = \frac{1}{\sqrt{s}_n} g(\frac{t-u_n}{s_n}) e^{i\xi_n t} \tag{3.21}$$

are chosen to best match the structures of f. This procedure approximates efficiently any signal structure that is well-localized in the time-frequency plane, regardless of whether this localization is in time or in frequency.

To any matching pursuit expansion[39][46], we can associate a time-frequency energy distribution defined by

$$Ef(t,\omega) = \sum_{n=0}^{\infty} |< R^n f, g_{\gamma_n} >|^2 W g_{\gamma_n}(t,\omega), \tag{3.22}$$

where

$$W g_{\gamma_n}(t,\omega) = 2\exp\left[-2\pi\left(\frac{(t-u)^2}{s^2} + s^2(\omega-\xi)^2\right)\right],$$

is the Wigner distribution [8] of the Gabor atom g_{γ_n}. Its energy is concentrated in the time and frequency domains where g_{γ_n} is localized. Figure 3.2 shows $Ef(t,\omega)$ for the signal f of 512 samples displayed in Fig. 3.1. This signal is built by adding waveforms of different time-frequency localizations. It is the sum of $\cos((1-\cos(ax))bx)$, two truncated sinusoids, two Dirac functions, and $\cos(cx)$. Each Gabor time-frequency atom selected by the matching pursuit is a dark elongated Gaussian blob in the time-frequency plane. The arch of the the $\cos((1-cos(ax))bx)$ is decomposed into a sum of atoms that covers its time-frequency support. The truncated sinusoids are in the center and upper left-hand corner of the plane. The middle horizontal dark line is an atom well localized frequency that corresponds to the component $cos(cx)$ of the signal. The two vertical dark lines are atoms very well localized in time that correspond to the two Diracs.

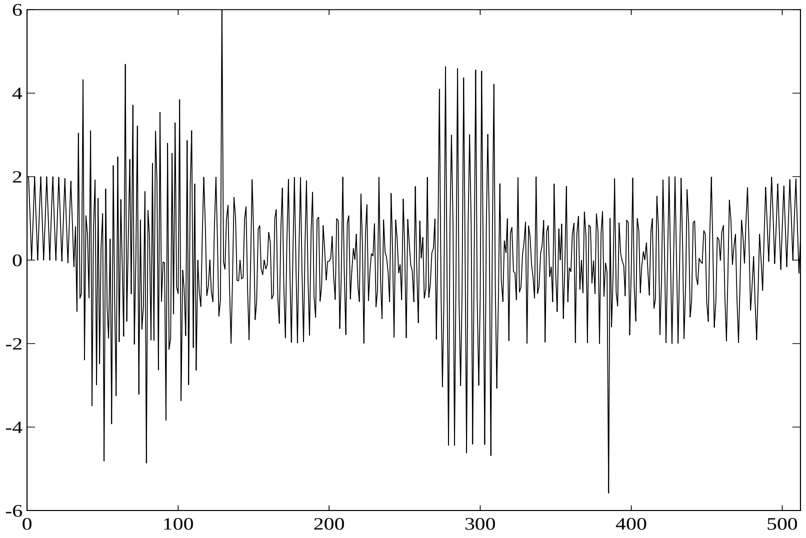

Figure 3.1: Synthetic signal of 512 samples built by adding $\cos((1-\cos(ax))bx)$, two truncated sinusoids, two Dirac functions, and $\cos(cx)$.

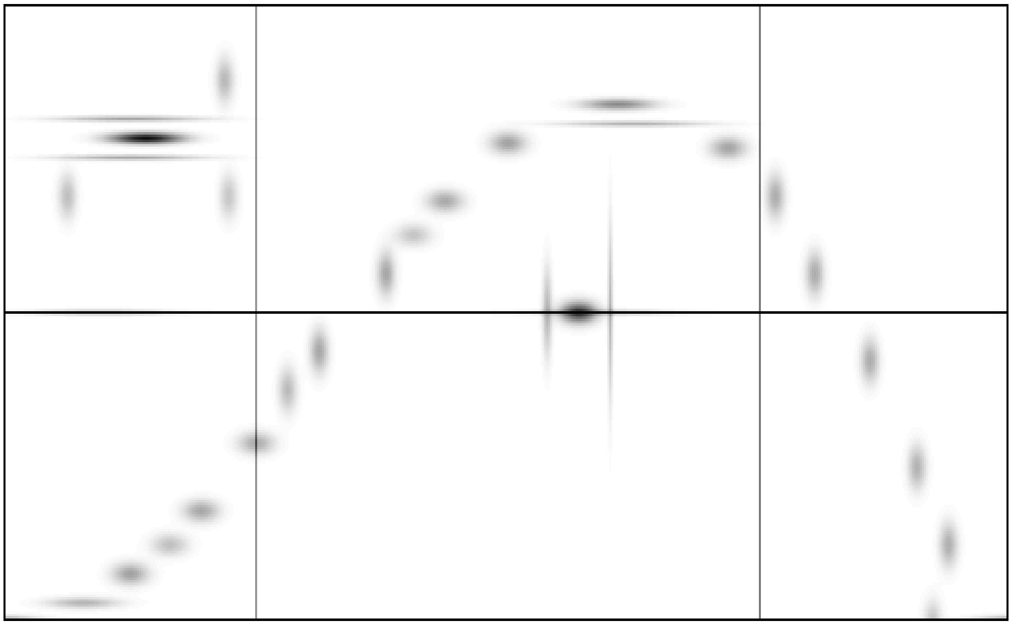

Figure 3.2: Time frequency energy distribution $Ef(t,\omega)$ of the signal in the figure above. The horizontal axis is time and the vertical axis is frequency. The darkness of the image increases with $Ef(t,\omega)$.

Chapter 4

Orthogonal Matching Pursuits

Matching pursuits do not in general converge in a finite number of iterations in a finite dimensional space. The reason is that successive iterations of the algorithm can reintroduce components into the residue of dictionary elements already removed. In section 4.1 we present an orthogonalized version of the matching pursuit algorithm which converges in a finite number of iterations in a finite dimensional space. The stability of the selected elements is of great importance for orthogonal pursuits. In section 4.2 we show that it is possible for both orthogonal and non-orthogonal pursuits to select degenerate collections of dictionary elements for function expansions, even if alternative, non-degenerate expansions exists. We describe a numerical implementation of an orthogonal pursuit and compare the computational complexity to that of a non-orthogonal matching pursuit in in section 4.3. In section 4.4 we compare the accuracy and stability of non-orthogonal and orthogonal pursuits with a dictionary of time-frequency atoms on a set of speech data.

4.1 Orthogonal Matching Pursuits

The approximations derived from a matching pursuit can be refined by orthogonalizing the directions of projection. The resulting orthogonal pursuit converges with a finite number of iterations in finite dimensional spaces, which is not the case for a non-orthogonal pursuit. A similar algorithm has been developed independently and in parallel by [43].

At each iteration, the vector g_{γ_k} selected by the matching algorithm is *a priori* not orthogonal to the previously selected vectors $\{g_{\gamma_p}\}_{0\leq p<k}$. In subtracting the projection of $R^k f$ over g_{γ_k} the algorithm reintroduces new components in the directions of $\{g_{\gamma_p}\}_{0\leq p<k}$. This can be avoided by orthogonalizing $\{g_{\gamma_p}\}_{0\leq p<k}$ with a Gram-Schmidt procedure. Let $u_0 = g_{\gamma_0}$. As in a matching pursuit, we choose g_{γ_k} that satisfies (3.4). This vector is orthogonalized with respect to the previously selected vectors by computing

$$u_k = g_{\gamma_k} - \sum_{p=0}^{k-1} \frac{< g_{\gamma_k}, u_p >}{||u_p||^2} u_p. \tag{4.1}$$

The residue is then defined by

$$R^{k+1}f = R^k f - \frac{< R^k f, u_k >}{||u_k||^2} u_k. \tag{4.2}$$

The vector $R^k f$ is the orthogonal projection of f on the orthogonal complement to the space generated by the vectors $\{g_{\gamma_p}\}_{0 \leq p < k}$. Equation (4.1) implies that $< R^k f, u_k > = < R^k f, g_{\gamma_k} >$ and thus

$$R^{k+1}f = R^k f - \frac{< R^k f, g_{\gamma_k} >}{||u_k||^2} u_k. \tag{4.3}$$

Since $R^{k+1}f$ and u_k are orthogonal,

$$||R^{k+1}f||^2 = ||R^k f||^2 - \frac{|< R^k f, g_{\gamma_k} >|^2}{||u_k||^2}. \tag{4.4}$$

If $R^k f \neq 0$, $< R^k f, g_{\gamma_k} > \neq 0$ and since $R^k f$ is orthogonal to all previously selected vectors the selected vectors $\{g_{\gamma_p}\}_{0 \leq p < k}$ are linearly independent. Since $R^0 f = f$, from equations (4.3) and (4.4), similarly to equations (3.7) and (3.8), we derive that for any $n > 0$

$$f = \sum_{0 \leq k < n} \frac{< R^k f, g_{\gamma_k} >}{||u_k||^2} u_k + R^n f, \tag{4.5}$$

and

$$\|f\|^2 = \sum_{0 \leq k < n} \frac{|< R^k f, g_{\gamma_k} >|^2}{||u_k||^2} + \|R^n f\|^2. \tag{4.6}$$

The theorem below proves that that the residues of an orthogonal pursuit converge strongly to zero and that the number of iterations required for convergence is less than or equal to the dimension of the space $\mathcal{H}$. Thus in finite dimensional spaces, orthogonal matching pursuits are guaranteed to converge in a finite number of steps, unlike non-orthogonal pursuits.

Theorem 4.1 *Let $\mathcal{H}$ be an N-dimensional Hilbert space and let $f \in \mathcal{H}$ (N may be infinite). An orthogonal pursuit converges in less than or equal to N iterations. The residue $R^n f$ defined in (4.3) satisfies*

$$\lim_{n \to N-1} ||R^n f|| = 0. \tag{4.7}$$

Hence

$$f = \sum_{n=0}^{N-1} \frac{< R^n f, g_{\gamma_n} >}{||u_k||^2} u_n \tag{4.8}$$

and

$$\|f\|^2 = \sum_{n=0}^{N-1} \frac{|< R^n f, g_{\gamma_n} >|^2}{||u_k||^2}. \tag{4.9}$$

Proof: We first suppose that $< R^k f, g_{\gamma_k} > \neq 0$ for $k < N$. If not, then we are through, since condition (3.3) implies that $< R^k f, g_\gamma >= 0$ for all $g_\gamma \in \mathcal{D}$, and because the vectors in $\mathcal{D}$ span $\mathcal{H}$, we must have $R^k f = 0$. Because $R^k f$ is orthogonal to $\{g_{\gamma_p}\}_{0 \leq p < k}$ and $< R^k f, g_{\gamma_k} > \neq 0$, the set $\{g_{\gamma_p}\}_{0 \leq p \leq k}$ must be linearly independent. When N is finite, the N linearly independent vectors $\{g_{\gamma_p}\}_{0 \leq p < N-1}$ form a basis of $\mathcal{H}$, and the orthogonalized vectors $\{u_p\}$ form an orthogonal basis of $\mathcal{H}$. The result follows directly.

When N is infinite, we have from the Bessel inequality that

$$\sum_{k=0}^{\infty} \frac{|< f, u_k >|^2}{||u_k||^2} \leq ||f||^2. \tag{4.10}$$

By (3.3), we must have

$$\lim_{k \to \infty} \sup_{\gamma \in \mathbf{\Gamma}} |< R^k f, g_\gamma >| = 0, \tag{4.11}$$

so $R^k f$ converges weakly to 0. To show strong convergence, we compute for $n < m$ the difference

$$\begin{aligned} ||R^n f - R^m f||^2 &= \sum_{k=n+1}^{m} \frac{|< f, u_k >|^2}{||u_k||^2} \\ &\leq \sum_{k=n+1}^{\infty} \frac{|< f, u_k >|^2}{||u_k||^2}, \end{aligned} \tag{4.12}$$

which goes to zero as n goes to infinity since the sum is bounded. The Cauchy criterion is satisfied, so $R^n f$ converges strongly to its weak limit of 0, thus proving the result.

□

The orthogonal pursuit yields a function expansion over an orthogonal family of vectors $\{u_k\}_{0 \leq p < n}$. To obtain an expansion of f over $\{g_{\gamma_n}\}_{0 \leq k < N}$ we must make a change of basis. The Gram-Schmidt vector u_k can be expanded within $\{g_{\gamma_p}\}_{0 \leq p \leq k}$

$$u_k = \sum_{p=0}^{n} b_{p,k} g_{\gamma_p}. \tag{4.13}$$

Inserting this expression into (4.8) yields

$$f = \sum_{n=0}^{M} \frac{< R^n f, g_{\gamma_n} >}{||u_n||^2} \sum_{p=0}^{n} b_{p,n} g_{\gamma_p}. \tag{4.14}$$

In the infinite dimensional case, without absolute convergence of the infinite series, we cannot rearrange the terms of this double summation to obtain

$$f = \sum_{0 \leq p < M} g_{\gamma_p} \sum_{p \leq n < M} b_{p,n} \frac{< R^n f, g_{\gamma_n} >}{||u_n||^2}. \tag{4.15}$$

The second summation that defines the expansion coefficients over the family $\{g_{\gamma_p}\}_{0\leq p<M}$ can indeed diverge. This happens when the family of selected elements is not a Riesz basis of the closed space it generates.

The residues of orthogonal matching pursuits in general decrease faster than the non-orthogonal matching pursuits. However, this orthogonal procedure can yield unstable expansions by selecting ill-conditioned family of vectors. It also requires many more operations to compute because of the Gram-Schmidt orthogonalization procedure. In the next section we show that in some cases, it is impossible to invert orthogonal pursuit function expansions, even when the dictionary contains a frame. The computational implementation and complexity of these two algorithms is compared in section 4.4.

4.2 Stability of Pursuits

The following theorem proves that there exist functions and dictionaries for which the orthogonal pursuit expansion in terms of the orthogonalized vectors u_k cannot be transformed into an expansion in terms of the original dictionary elements g_γ, even when the dictionary contains an orthonormal basis. We construct a function which has energy spread all across the orthonormal basis. The ill-conditioned elements are selected because they are much better suided to expanding the function in question than the vectors of the orthonormal basis.

Theorem 4.2 *There exists a function $f \in L^2[0,1]$ and a dictionary for $L^2[0,1]$ that contains an orthonormal basis such that the dictionary elements in the orthogonal matching pursuit expansion of f do not form a Riesz basis.*

Proof: We first define a set of functions on $L^2[0,1]$ which we will use to construct our dictionary and our function. Let

$$w_k(t) = 2^{2k}\chi_{[1-2^{-k},1-2^{-k}+2^{-4k}]}(t), \tag{4.16}$$

for $k \geq 1$. The w_k's are a set of step functions which tend towards a delta function at $t=1$ as k goes to infinity. Our dictionary consists of the Fourier basis for $L^2[0,1]$ together with the set $(x_k)_{k\geq 1}$ defined by,

$$\begin{aligned} x_1(t) &= w_1(t) && (4.17)\\ x_k(t) &= (1-2^{-k})^{\frac{1}{2}}w_1(t) + 2^{-\frac{k}{2}}w_k(t). && (4.18)\end{aligned}$$

We have

$$\mathcal{D} = (x_k)_{k\geq 1} \cup (e^{2\pi ikt})_{k\in\mathbf{Z}}. \tag{4.19}$$

We take as our function to decompose,

$$f = \sum_{k=1}^{\infty} 2^{-2k} w_k. \tag{4.20}$$

The Fourier basis is ill-suited to representing highly spatially localized functions such as the w_k's which comprise f. Hence, the pursuit will select the localized functions x_k for the expansion, despite the degeneracy of this set.

We prove by induction that the selected dictionary elements are $x_1, x_2, \ldots$, and that $R^n f$ is given by

$$R^n f = \sum_{k=n+1}^{\infty} 2^{-2k} w_k. \tag{4.21}$$

The expression holds for $R^0 f = f$. For f we have,

$$\begin{aligned} |< f, e^{2\pi imt} >| &= |\int_0^1 R^n f(x) e^{-2\pi imt} dx| && (4.22) \\ &\leq \int_0^1 |R^n f| && (4.23) \\ &= \sum_{k=1}^{\infty} 2^{-2k} 2^{2k} 2^{-4k} = \frac{1}{15} && (4.24) \end{aligned}$$

and

$$\begin{aligned} |< f, x_1 >| &= 2^{-2} && (4.25) \\ |< f, x_m >| &= (1-2^{-m})^{\frac{1}{2}} 2^{-2} + 2^{-\frac{m}{2}} 2^{-2m} < |< f, x_1 >|. && (4.26) \end{aligned}$$

Hence the first selected dictionary element will be x_1.

We now suppose that $R^n f = \sum_{k=n+1}^{\infty} 2^{-2k} w_k$. We compute the inner product of $R^n f$ with all elements of the dictionary to determine the next selected element. For the Fourier basis, we have

$$\begin{aligned} |< R^n f, e^{2\pi imt} >| &\leq \sum_{k=n+1}^{\infty} 2^{-2k} 2^{2k} 2^{-4k} && (4.27) \\ &= \frac{2^{-4n}}{15}. && (4.28) \end{aligned}$$

For the v_k's we have

$$\begin{aligned} |< R^n f, x_m >| &= 0, \text{ for } m \leq n && (4.29) \\ &= 2^{-\frac{5}{2}m}, \text{ for } m > n, && (4.30) \end{aligned}$$

so we have $\max_{g_\gamma \in \mathcal{D}} |< R^n f, g_\gamma >| = |< R^n f, x_{n+1} >|$. The normalized projection of x_{n+1} onto the complement of the span of the set $\{x_1, \ldots x_n\}$ is w_{n+1}. We remove this normalized projection from $R^n f$ to obtain,

$$R^{n+1} f = \sum_{k=n+2}^{\infty} 2^{-2k} w_k. \tag{4.31}$$

The set $(x_k)_{k\geq 1}$ is clearly not a Riesz basis, because

$$\sum_{k=1}^{\infty} |< x_k, x_1 >|^2 = \infty \tag{4.32}$$

and

$$\lim_{n\to\infty} \sum_{k=1}^{\infty} |< x_k, \frac{x_n - < x_n, x_1 > x_1}{\|x_n - < x_n, x_1 > x_1\|} >|^2 \tag{4.33}$$

$$= \lim_{n\to\infty} \sum_{k=1}^{\infty} |< x_k, w_n >|^2 = 0. \tag{4.34}$$

□

An equivalent result holds for non-orthogonal matching pursuits but the proof is a bit more complicated. Because non-orthogonal pursuits yield expansions in terms of dictionary elements, rather than in terms of the orthogonalized dictionary elements, the selection of a degenerate collection does not pose difficulties.

Theorem 4.3 *There exists a function $f \in L^2[0,1]$ and a dictionary for $L^2[0,1]$ that contains an orthonormal basis such that the dictionary elements in the non-orthogonal matching pursuit expansion of f do not form a Riesz basis.*

We show that functions and dictionaries exist for which the selected dictionary elements fail to satisfy the upper or lower frame bounds. We first construct from the functions w_k above dictionary a function for which the selected elements do not satisfy the upper frame bound.

Our dictionary consists of the Fourier basis for $L^2[0,1]$ together with the degenerate sets $(x_k)_{k\geq 1}$ and $(y_k)_{k\geq 1}$ defined by,

$$x_1 = y_1 = w_1 \tag{4.35}$$

$$x_k = \frac{1}{\sqrt{2}} w_1 + \frac{1}{\sqrt{2}} w_k \tag{4.36}$$

$$y_k = \frac{1}{\sqrt{2}} w_1 - \frac{1}{\sqrt{2}} w_k \tag{4.37}$$

We have

$$\mathcal{D} = (x_k)_{k\geq 1} \cup (y_k)_{k\geq 1} \cup (e^{2\pi ikt})_{k\in\mathbf{Z}}. \tag{4.38}$$

We again take as our function to decompose,

$$f = \sum_{k=1}^{\infty} 2^{-2k} w_k. \tag{4.39}$$

The idea of the proof is the same as before. The additional elements in the dictionary are necessary because when the non-orthogonalized pursuit removes x_k from the current residue, it introduces a component of w_1 into the next residue. We remove y_k next in order to eliminate this w_1 component and keep the problem compartmentalized.

We prove by induction that the residue $R^{2n+1}f = \sum_{k=n+1}^{\infty} 2^{-2k} w_k$. We assume that we have a choice function which selects the x_k's over the y_k's when maximum inner products are equal. The first selected dictionary elements will be x_1, followed by $x_2, y_2, x_3, y_3, \ldots$.

We first show that $R^1 f = \sum_{k=1}^{\infty} 2^{-2k} w_k$. We have

$$\begin{aligned} |< f, x_1 >| &= |< f, y_1 >| = 2^{-2}, & (4.40)\\ |< f, x_m >| &= \frac{1}{\sqrt{2}} 2^{-2} + \frac{1}{\sqrt{2}} 2^{-2m}, \text{ for } m > 1 & (4.41)\\ |< f, y_m >| &= \frac{1}{\sqrt{2}} 2^{-2} - \frac{1}{\sqrt{2}} 2^{-2m}, \text{ for } m > 1 & (4.42)\\ & & (4.43)\end{aligned}$$

and

$$|< f, e^{2\pi imt} >| \leq \frac{1}{15}, \qquad (4.44)$$

so the first selected dictionary element is x_1. We thus obtain $R^1 f = \sum_{k=2}^{\infty} 2^{-2k} w_k$.

Suppose now that we have $R^{2n+1}f = \sum_{k=n+1}^{\infty} 2^{-2k} w_k$. The inner products of $R^{2n+1}f$ with the dictionary elements will be

$$\begin{aligned} |< R^{2n+1}f, x_m >| &= |< R^{2n+1}f, y_m >| = 0, \text{ for } m \leq n & (4.45)\\ |< R^{2n+1}f, x_m >| &= |< R^{2n+1}f, y_m >| = \frac{1}{\sqrt{2}} 2^{-2m}, \text{ for } m > n & (4.46)\\ |< R^{2n+1}f, e^{2\pi imt} >| &\leq \frac{2^{-4n}}{15}. & (4.47)\end{aligned}$$

Thus, the next selected element will be x_{n+1} and

$$R^{2n+2}f = -\frac{1}{2}(w_1 - w_{n+1}) + \sum_{k=n+2}^{\infty} 2^{-2k} w_k. \qquad (4.48)$$

The inner products of $R^{2n+2}f$ with the dictionary elements will be

$$\begin{aligned} |< R^{2n+2}f, x_m >| &= 0, \text{ for } m \leq n+1 & (4.49)\\ |< R^{2n+2}f, x_m >| &= \frac{1}{\sqrt{2}} 2^{-2m}, \text{ for } m > n+1 & (4.50)\\ |< R^{2n+2}f, y_m >| &= 0, \text{ for } m \leq n & (4.51)\\ |< R^{2n+2}f, y_m >| &= \frac{1}{\sqrt{2}} 2^{-2m}, \text{ for } m > n & (4.52)\\ |< R^{2n+2}f, e^{2\pi imt} >| &\leq 2^{-2(n+1)}[-\frac{1}{2} 2^{-4} + \frac{1}{2} 2^{-4(n+1)}] + 2^{-4(n+2)} \frac{16}{15}. & (4.53)\end{aligned}$$

The maximum inner product will be $| < R^{2n+2}f, y_{n+1} > |$. Subtracting the component of y_{n+1} from $R^{2n+2}f$, we obtain the desired result, $R^{2(n+1)+1}f = \sum_{k=n+2}^{\infty} 2^{-2k} w_k$.

We have

$$\sum_{k=1}^{\infty} | < x_m, x_k > |^2 = \sum_{k=1}^{\infty} | < x_m, y_k > |^2 = \tag{4.54}$$

$$\sum_{k=1}^{\infty} | < y_m, x_k > |^2 = \sum_{k=1}^{\infty} | < y_m, y_k > |^2 = \infty, \tag{4.55}$$

so we see that the set of selected dictionary elements is degenerate.

□

We now construct a dictionary and function for which the selected elements do not satisfy the lower frame bound. We create three orthogonal families of step functions from which we will build our dictionary and our function. We let

$$x_k(t) = 2^{2(k+4)} \chi_{[1-2^{-k}, 1-2^{-k}+2^{-4(k+4)}]}(t) \tag{4.56}$$

$$y_k(t) = 2^{2(k+4)} \chi_{[1-2^{-k}+2^{-4(k+4)}, 1-2^{-k}+2\ 2^{-4(k+4)}]}(t) \tag{4.57}$$

$$z_k(t) = 2^{2(k+4)} \chi_{[1-2^{-k}+22^{-4(k+4)}, 1-2^{-k}+3\ 2^{-4(k+4)}]}(t). \tag{4.58}$$

Our dictionary will consist of the Fourier basis $(e^{2\pi int})_{n\in\mathbf{Z}}$ together with three orthonormal families (x_k), (y_k), and (z_k). We will build our function f from a sum of orthogonal pieces, piece k of which will be decomposed into the sequence x_k, y_k, z_k.

The three families which form the dictionary are

$$a_k = \frac{3}{5} x_k + \frac{4}{5} y_k \tag{4.59}$$

$$b_k = x_k \tag{4.60}$$

$$c_k = \sqrt{1-\epsilon_k^2}\, y_k + \epsilon_k z_k. \tag{4.61}$$

Here (ϵ_k) is a sequence which decreases monotonically to 0. Our function f will consist of a linear sum of the functions

$$f_k = 4x_k + 3y_k - \frac{21}{25} \frac{\epsilon_k}{\sqrt{1-\epsilon_k^2}}. \tag{4.62}$$

We use the cycle of 3 elements, so that we can progressively eliminate the pieces f_k from f without any of the selected elements having to contain a large component of z_k. We also must build the 3-cycle so that z_k is contained in its span. We can then show that the selected elements are not a frame because the sum of the squares of their inner products with z_k can be made arbitrarily small.

By adjusting the heights and widths of the functions $(x_k), (y_k)$, and (z_k), we can make the magnitude of the inner products of the residues with the Fourier basis as small as we like. By adjusting the rate of decay of the coefficients of the f_k we can ensure that a_k is selected, followed by b_k then c_k. We take

$$f = \sum_{k=1}^{\infty} 2^{-3k} f_k. \tag{4.63}$$

The details of the proof are similar to the cases above.

□

4.3 Numerical Implementation of Orthogonal Matching Pursuits

We suppose that $\mathcal{H}$ is a finite dimensional space and $\mathcal{D}$ a dictionary with a finite number of vectors. The optimality factor α is set to 1.

The initialization and selection portions of the orthogonal matching pursuit algorithm are implemented in the same way as they are for the non-orthogonal algorithm. The difference between the two algorithms is in the updating of the inner products $< R^n f, g_\gamma >$ after a vector has been selected. Once the vector g_{γ_n} is selected, we must compute the expansion coefficients of the orthogonal vector u_n

$$u_n = \sum_{p=0}^{n} b_{p,n} g_{\gamma_p}. \tag{4.64}$$

The Gram-Schmidt orthogonalization scheme can be used to obtain these coefficients in $O(n^2 I)$ time, but the numerical properties of this scheme are not good. A better method is to use the fact that

$$u_n = g_{\gamma_n} - P_{n-1} g_{\gamma_n} \tag{4.65}$$

where P_{n-1} is the projection onto the span of the vectors $\{g_{\gamma_k}\}_{0 \leq k \leq n-1}$. We can write this projection operator P_{n-1} as the product

$$P_{n-1} = G_{n-1}(G_{n-1}^* G_{n-1})^{-1} G_{n-1}^*, \tag{4.66}$$

where G_{n-1} is an n by n matrix which has as its columns the vectors $\{g_{g_{\gamma_k}}\}_{0 \leq k \leq n-1}$. For $p < n$ the coefficient $b_{p,n}$ is given by the p^{th} entry in the column vector $-(G_{n-1}^* G_{n-1})^{-1} G_{n-1}^* g_{\gamma_n}$, and for $p = n$ we have $b_{n,n} = 1$. We compute the $b_{p,n}$'s in two steps. We first form the vector $G_{n-1}^* g_{\gamma_n}$. This requires computing the n inner products $< g_{\gamma_k}, g_{\gamma_n} >$ for $0 \leq k < n$ and so requires $O(nI)$ operations. We then form the matrix $G_{n-1}^* G_{n-1}$. This matrix is built recursively by adding an additional row and column to $G_{n-2}^* G_{n-2}$. This row and column have entries of the form $< g_{\gamma_k}, g_{\gamma_{n-1}} >$ for $k \leq n-1$, and since we compute these values in step $n-1$ while forming of $G^*_{n-2}g_{\gamma_{n-1}}$, this step requires no additional computations. We can use the LDU decomposition from step $n-1$ as part of a block LDU decomposition [25], so computation of the inverse requires $O(n^2)$ steps. The total work required to compute the coefficients $b_{p,n}$ is then $O(n^2)$.

We then compute the inner product of the new residue $R^{n+1}f$ with all $g_\gamma \in \mathcal{D}_\alpha$ using the orthogonal updating formula (4.3)

$$< R^{n+1}f, g_\gamma >=< R^n f, g_\gamma > - < R^n f, g_{\gamma_n} > < u_n, g_\gamma > . \tag{4.67}$$

Since

$$< u_n, g_\gamma >= \sum_{p=0}^{n} b_{p,n} < g_{\gamma_p}, g_\gamma >, \tag{4.68}$$

computing $\{< R^{n+1}f, g_\gamma >\}_{\gamma\in\mathbf{\Gamma}}$ requires $O(nIZ)$ operations. The total number of operations to compute P orthogonal matching pursuit iterations is therefore $O(P^3+P^2IZ)$. For P iterations, the non-orthogonal pursuit algorithm is P times faster than the orthogonal one. When P is large, which is the case in many signal processing applications, the orthogonal pursuit algorithm is much slower and requires too many calculations for real time processing. When P remains small, the orthogonal pursuit is more advantageous because it converges faster.

4.4 Comparison of Non-orthogonal and Orthogonal Pursuits

4.4.1 Accuracy

To compare the performance of the orthogonal and non-orthogonal pursuits, we segmented a digitized speech recording into 512-sample pieces and decomposed the pieces using both algorithms. The dictionary used was the discretized described in section 3.3.

Figure 4.1 shows for both algorithms the decay of the residual $||R^n f||$ as a function of n for a 512 sample speech segment. When n is close to 512, the dimension of the signal space, the orthogonal pursuit residue converges very rapidly to 0. The non-orthogonal pursuit, on the other hand, converges exponentially with a slow rate when n is large. We see, then, that orthogonal pursuits yield much better approximations when n is large.

The performance of the two algorithms is similar in the early part of the expansion, however. The reason is that for the early part of the expansion the selected vectors are nearly orthogonal, so the orthogonalization step does not contribute greatly. This near-orthogonality comes from the fact that for both pursuits $< R^{n+1}f, g_{\gamma_n} >= 0$, so

$$\frac{|< R^{n+1}f, g_\gamma >|^2}{||R^{n+1}f||^2} \leq (1 - |< g_\gamma, g_{\gamma_n} >|^2). \tag{4.69}$$

The vector $g_{\gamma_{n+1}}$ is chosen by finding the $\gamma \in \Gamma$ for which the left hand side of (4.69) is maximized. We see from (4.69) that there is a penalty for selecting dictionary elements g_γ

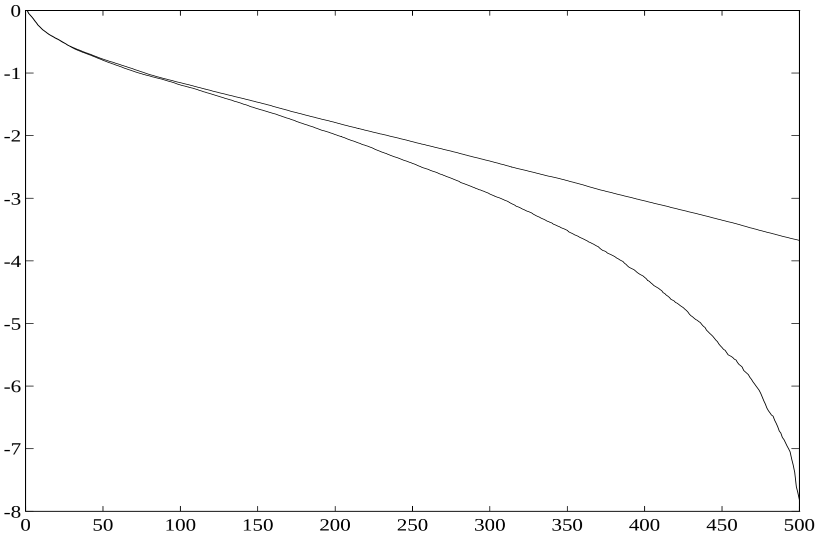

Figure 4.1: $\log \|R^n f\|$ as a function of the n. The top curve shows the decay of $\|R^n f\|$ for a non-orthogonal pursuit and the bottom for an orthogonal pursuit.

for which $|< g_\gamma, g_{\gamma_n} >|$ is large. Provided that $|< g_{\gamma_n}, g_{\gamma_{n+1}} >|$ is small, a similar (but smaller) penalty exists for selecting a $g_{\gamma_{n+2}}$ which correlates with either g_{γ_n} or $g_{\gamma_{n+1}}$, and so on. Hence, the initially selected vectors tend to be orthogonal.

These nearly-orthogonal elements which comprise the initial terms of the expansion correspond to the signal's "coherent structures," the portions of the signal which are well-approximated by dictionary elements. We describe these coherent structures in more detail in chapter 7. The correlation ratio, defined by

$$\lambda(R^n f) = \sup_{\gamma \in \mathbf{\Gamma}} \frac{|< R^n f, g_\gamma >|}{\|R^n f\|} \tag{4.70}$$

is an important measure of the degree to which structures in the residue $R^n f$ resemble dictionary elements. A signal f which possess structures which are well-represented by dictionary elements will have large values of $\lambda(f)$. As the matching pursuit proceeds, these structures are removed, and $\lambda(R^n f)$ decreases. Experiments show that $\lambda(R^n f)$ converges to a dictionary dependent constant, λ_∞. The coherent structures of f are defined to be those structures selected before $\lambda(R^n f)$ is sufficiently close to the value λ_∞. We denote by $N_c(f)$ the number of coherent structures in f.

For many applications, we are interested in only the coherent portion of the expansion of f. Although for large expansions, the orthogonal pursuit produces a much smaller error, for the coherent portion of the expansion, the difference between the two algorithms is not great. For the discretized Gabor dictionary with 512 samples, we find that $\lambda_\infty \approx 0.17$. Selected dictionary elements are deemed to be coherent until a running average of the $||R^n f||$'s is within 2 percent of λ_∞.

For the 274 speech segments tested, the average number of coherent structures was 72.7. For the coherent portion of the signal, the norm of the residual generated by the orthogonal pursuit was on average only 18.5 percent smaller than the norm of the residual for the matching pursuit. More precisely, let $R^n f$ denote the non-orthogonal pursuit residue, and let $R^n_o f$ denote the orthogonal pursuit residue. For the speech segments tested, the ratio $\frac{||R^{N_c} f||}{||R_o^{N_c} f||}$ ranged from 0.864 to 1.771 with an average of 1.185 and a standard deviation of 0.176. We see, then, that for the coherent part of the signal, the benefits of the orthogonalization are not large.

The computational cost of performing a given number of iterations of an orthogonal pursuit is much higher than for the non-orthogonal pursuit, as we showed in the last section. However, because of the better convergence properties of the orthogonal pursuit, we need not perform as many iterations to obtain the same accuracy as a non-orthogonal pursuit. For the coherent portions of the tested speech segments, the orthogonal pursuit required an average of $N_c - 4$ iterations to obtain an error equivalent to that of the non-orthogonal pursuit with N_c iterations. The implementation of the pursuit used requires $I = O(1)$ operations to compute the inner products $< g_\gamma, g_{\gamma'} >$ and on average, $Z = N = 512$ of these inner products are non-zero. The non-orthogonal expansion of the coherent part of the signal thus requires roughly $4 \times 10^4 I$ operations whereas the orthogonal expansion requires roughly $2 \times 10^6 I$ operations. The cost is two orders of magnitude higher for a 20 percent improvement in the error.

4.4.2 Stability

Orthogonal pursuits yield expansions of the form

$$f = \sum_{k=0}^{n} \beta_k u_k + R^n f \tag{4.71}$$

where the u_k's are orthogonalized dictionary elements. When the selected set of dictionary elements is degenerate (when the set does not form a Riesz basis for the space it spans), these expansions cannot be converted into expansions over the dictionary elements g_γ. Our results from section 4.2 shows that this is a legitimate concern, at least in theory. We now examine numerically the stability of the collection of dictionary elements selected by orthogonal and non-orthogonal pursuits.

To compare the degeneracy of the sets of elements selected by the two algorithms, we computed the 2-norm condition number for the Gram matrix $G_{i,j} =< g_{\gamma_i}, g_{\gamma_j} >$ for twenty 128-sample speech segments. Figure 4.2 shows for both pursuits the condition number $\kappa(n)$ of the Gram matrix as a function of the number of iterations n for one 128-sample speech segment. As we discussed above, the initially selected coherent structures are roughly

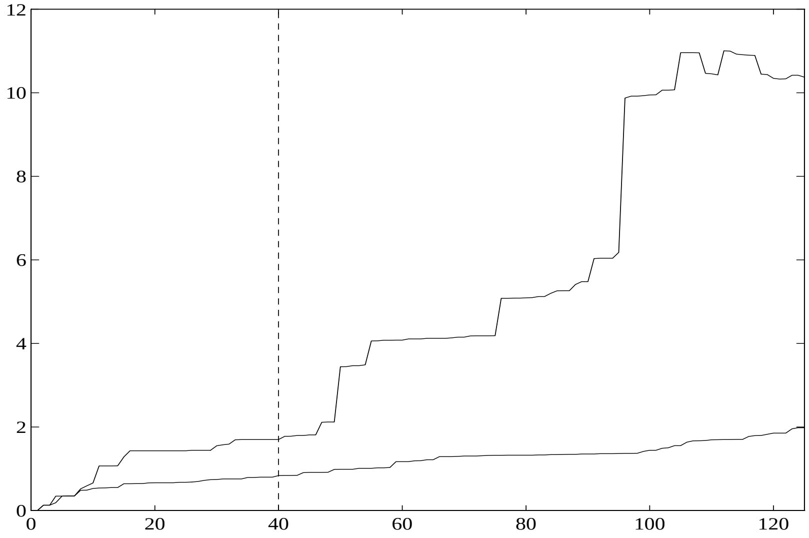

Figure 4.2: $\log_{10} \kappa(n)$ for the Gram matrix of the selected dictionary elements as a function of the number of iterations. The top curve is the condition number for a non-orthogonal pursuit, and the bottom is for an orthogonal pursuit. The dashed line is at N_c.

orthogonal and form a well-conditioned set for both pursuits. As the pursuit proceeds, the set selected by the non-orthogonal pursuit grows more and more singular, while the set selected by the orthogonal pursuit stays well-conditioned. The reason is that for the non-orthogonal pursuit, the penalty (4.69) against selecting a $g_{\gamma_{n+k}}$ that correlates with g_{γ_n} decreases as k increases. Hence, as the number of iterations increases beyond the number of coherent structures, the set grows more and more singular. For the orthogonal pursuit, on the other hand, we have

$$\frac{|< R^{n+1}f, g_\gamma >|^2}{||R^{n+1}f||^2} \leq ||(I - P_n)g_\gamma||^2, \tag{4.72}$$

where P_n is the orthogonal projection onto the space spanned by $g_{\gamma_0} \ldots g_{\gamma_n}$. Hence there is a penalty against selecting for $g_{\gamma_{n+1}}$ a g_γ which correlates strongly with *any* of the previously selected elements. The table below summarizes the results.

Pursuit	mean(log($\kappa(N)$))	mean(log($\kappa(N_c)$))
Non-orthogonal	12.2	1.53
Orthogonal	2.09	0.621

Chapter 5

Group Invariant Dictionaries

The translation, dilation, and frequency modulation of any vector that belongs to the Gabor dictionary still belongs to this dictionary. The dictionary invariance under the action of any operators that belong to the group of translations, dilations or frequency modulations implies important invariance properties of the matching pursuit. We study such properties when the dictionary is left invariant by any given group of unitary linear operators $\mathcal{G} = \{G_\tau\}_{\tau\in\Omega}$ that is a representation of the group Ω. Since each operator G_τ is unitary, its inverse and ajoint is $G_{\tau^{-1}}$, where τ^{-1} is the inverse of τ in Ω. For example, the unitary groups of translation, frequency modulation and dilation over $\mathcal{H} = \mathbf{L}^2(\mathbf{R})$ are defined respectively by $G_\tau f(t) = f(t-\tau)$, $G_\tau f(t) = e^{i\tau t} f(t)$ and $G_\tau f(t) = \frac{1}{\sqrt{s^\tau}} f(\frac{t}{s^\tau})$.

Definition 5.1 *A dictionary $\mathcal{D} = \{g_\gamma\}_{\gamma\in\mathbf{\Gamma}}$ is invariant with respect to the group of unitary operators $\mathcal{G} = \{G_\tau\}_{\tau\in\Omega}$ if and only if for any $g_\gamma \in \mathcal{D}$ and $\tau \in \Omega$, $G_\tau g_\gamma \in \mathcal{D}$.*

The Gabor dictionary is invariant under the group generated by the groups of translations, modulations and dilations. The properties of the corresponding matching pursuit depends upon the choice function C that chooses for any $f \in \mathcal{H}$ an element $g_{\gamma_0} = C(\mathbf{E}[f])$ from the set

$$\mathbf{E}[f] = \{g \in \mathcal{D} : | < f, g > | \geq \alpha \sup_{g_\gamma \in \mathcal{D}} | < f, g_\gamma > |\}$$

onto which f is then projected. The following proposition imposes a commutatitivity condition on the choice function C so that the matching pursuit commutes with the group operators.

Proposition 5.1 *Let $\mathcal{D}$ be invariant with respect to the group of unitary operators $\mathcal{G} = \{G_\tau\}_{\tau\in\Omega}$. Let $f \in \mathcal{H}$ and*

$$f = \sum_{k=0}^{n-1} a_n g_{\gamma_n} + R^n f$$

be its matching pursuit computed with the choice function C. If for any $n \in \mathbf{N}$

$$C G_\tau \mathbf{E}[R^n f] = G_\tau C \mathbf{E}[R^n f], \tag{5.1}$$

then the matching pursuit decomposition of $G_\tau f$ is

$$G_\tau f = \sum_{k=0}^{n-1} a_n G_\tau g_{\gamma_n} + G_\tau R^n f. \tag{5.2}$$

The condition (5.1) means that an element chosen from the $\mathbf{E}[R^n f]$ transformed by G_τ is the transformation by G_τ of the element chosen among $\mathbf{E}[R^n f]$. Equation (5.2) proves the vectors selected by the matching pursuit of $G_\tau f$ are the vectors selected for the matching pursuit of f transformed by G_τ and the residues of $G_\tau f$ are equal to the residues of f transformed by G_τ.

Proof: Since the group is unitary, for any $g_\gamma \in \mathcal{D}$

$$< G_\tau f, g_\gamma > = < f, G_{\tau^{-1}} g_\gamma > .$$

Hence $g_\gamma \in \mathbf{E}[G_\tau f]$ if and only if $G_{\tau^{-1}} g_\gamma \in \mathbf{E}[f]$ which proves that $\mathbf{E}[G_\tau f] = G_\tau \mathbf{E}[f]$. By using the commutativity (5.1) of the choice function with respect to G_τ we then easily prove (5.2) by induction.

□

The difficulty is now to prove that there exists a choice function that satisfies the commutativity relation (5.1) for all $f \in \mathcal{H}$ or at least for almost all $f \in \mathcal{H}$. The following proposition gives a necessary condition to construct such choice functions.

Proposition 5.2 *Let* $\mathbf{K}$ *be set of functions* $f \in \mathcal{H}$ *such that there exists* $G_\tau \neq I$ *with*

$$\mathbf{E}[f] = \mathbf{E}[G_\tau f]. \tag{5.3}$$

There exists a choice function C *such that for any* $f \in \mathcal{H} - \mathbf{K}$ *and* $G_\tau \in \mathcal{G}$

$$C G_\tau \mathbf{E}[f] = G_\tau C \mathbf{E}[f]. \tag{5.4}$$

Proof: To define such a choice function we construct the equivalence classes of the equivalence relation R_1 in $\mathcal{H}$ defined by $f\ R_1\ h$ if and only if there exists $G_\tau \in \mathcal{G}$ such that $f = G_\tau h$. The axiom of choice guarantees that there exists a choice function that chooses an element from each equivalence class. Let $\mathcal{H}_1$ be the set of all class representatives. The axiom of choice also guarantees that for any $f \in \mathcal{H}_1$ there exists a choice function C that associates to any set $\mathbf{E}[f]$ an element within this set. To extend this choice function we define a new equivalence relation R_2 in $\mathcal{H}$ defined by $f\ R_2\ h$ if and only if there exists $G_\tau \in \mathcal{G}$ such that $f = G_\tau h$ and $\mathbf{E}[f] = \mathbf{E}[h]$. All elements that belong to $\mathcal{H} - \mathbf{K}$ correpond to equivalence classes of 1 vector. For each equivalence class, we choose a representative f. There exists G_τ such that $G_\tau f \in \mathcal{H}_1$. Since G_τ is unitary, for any $g_\gamma \in \mathcal{D}$

$$< G_\tau f, g_\gamma > = < f, G_{\tau^{-1}} g_\gamma > .$$

Hence, $\mathbf{E}[G_\tau f] = G_\tau E[f]$. For any h that is equivalent to f with respect to R_2, we define

$$C(\mathbf{E}[h]) = C(\mathbf{E}[f]) = G_{\tau^{-1}} C(\mathbf{E}[G_\tau f]) = G_{\tau^{-1}} C(G_\tau \mathbf{E}[f]) \in E[f]. \tag{5.5}$$

This choice function associates a unique element to each different set $E[f]$. If $f \in \mathcal{H} - \mathbf{K}$, property (5.5) implies that this choice function satisfies (5.4).

□

Proposition 5.3 *When $\mathcal{H}$ is an infinite dimensional space, if for any $g_\gamma \in \mathcal{D}$ and $G_\tau \neq I$ there exists A such that for any $h \in \mathcal{H}$*

$$||h||^2 \geq A \sum_{n \in \mathbf{N}} | < h, G_{\tau^n} g_\gamma > |^2, \tag{5.6}$$

then $\mathbf{K} = \{0\}$.

Proof: If there exists $f \in \mathcal{H}$ and G_τ such that $\mathbf{E}[f] = \mathbf{E}[G_\tau f]$, then for any $n \in \mathbf{N}$, $\mathbf{E}[f] = \mathbf{E}[G_{\tau^n} f]$, where $G_{\tau^n} f$ is the n^{th} power of G_τ. Hence, for any $g_\gamma \in E[f]$ and $n \in \mathbf{N}$

$$| < f, G_{\tau^n} g_\gamma > | \geq \alpha \lambda(f).$$

If we set $h = f$ in (5.6), this property implies that $\lambda(f) = 0$ otherwise f would have an infinite norm. Since linear combinations of elements in $\mathcal{D}$ are dense in $\mathcal{H}$, if $\lambda(f) = 0$ then $f = 0$.

□

Property (5.6) is satisfied for the Gabor dictionary and the group $\mathcal{G}$ composed of dilations, translations and modulations for $\mathcal{H} = \mathbf{L}^2(\mathbf{R})$. This comes from our ability to construct frames of $\mathbf{L}^2(\mathbf{R})$ through translations and dilations or frequency modulations of Gaussian functions [13]. This result implies that there exists a choice function such that the matching pursuit in a Gabor dictionary commutes with dilations translations and frequency modulations.

For cyclic groups, we can derive necessary and sufficient conditions for the existence of a choice function such that $G_\tau R = R G_\tau$ in a complex, finite dimensional space. Examples of cyclic groups include unit translations modulo N and unit modulations. The following theorem shows that with a suitable dictionary we can always obtain translation invariant decompositions or modulation invariant decompositions.

Proposition 5.4 *Let $\mathcal{H}$ be a finite dimensional space. We can construct a choice function C for which $G_\tau R^n f = R^n G_\tau f$ if and only if the dictionary $\mathcal{D}$ contains all the eigenvectors of G_τ. Moreover, the constructed choice function has an optimality factor of 1 except on a set of measure 0.*

Proof: Suppose $\mathbf{E}[f] = \mathbf{E}[G_\tau f]$. We require that $RG_\tau f = G_\tau Rf$, so from $Rf = f - < f, C(\mathbf{E}[f]) > C(\mathbf{E}[f])$ we derive that

$$< G_\tau f, C(\mathbf{E}[G_\tau f]) > C(\mathbf{E}[G_\tau f]) =< f, C(\mathbf{E}[f]) > G_\tau C(\mathbf{E}[f]), \tag{5.7}$$

so we have

$$G_\tau C(\mathbf{E}[f]) = \beta C(\mathbf{E}[f]). \tag{5.8}$$

Plugging (5.8) into (5.7) gives $|\beta|^2 = 1$. Since G_τ is unitary, setting $C(\mathbf{E}[f])$ equal to any eigenvector of G_τ will satisfy (5.7), and these are the only solutions. Thus, we can satisfy $RG_\tau f = G_\tau Rf$ if and only if $C(\mathbf{E}[f])$ is an eigenvector of G_τ.

Suppose the dictionary $\mathcal{D}$ contains all the eigenvectors of G_τ. We can construct in two steps a choice function which for almost all f yields a vector g_{γ_0} such that $|< f, g_{\gamma_0} >| = \max_\gamma |< f, g_\gamma >|$ for almost all f, and a vector g_{γ_0} such that $|< f, g_{\gamma_0} >| \geq \frac{1}{\sqrt{N}} \max_\gamma |< f, g_\gamma >|$ for f in a set of measure 0.

On the set $\mathcal{H} - \mathbf{K}$ we construct a C_1 with optimality factor $\alpha = 1$ using the method of proposition 5.2. For f to be in the set $\mathbf{K}$, it must satisfy

$$\lambda(f) = | < f, g_\gamma > | = | < f, G_{\tau^{-1}} g_\gamma > |.$$

This set $\mathbf{K}$ is of measure 0 in $\mathcal{H}$. Thus, we are constructing a choice factor with an optimality factor of 1 except on a set of measure 0.

On the set $\mathbf{K}$ we compute $\mathbf{E}'[f]$ with optimality factor $\alpha = \frac{1}{\sqrt{N}}$. We partition $\mathbf{K}$ into the sets

$$\begin{aligned} \mathbf{K}_K &= \{f \in \mathbf{K} : \mathbf{E}'[f] = E'[G_\tau f] \text{ for } G_\tau \neq I\} \\ \mathbf{K}_H &= \mathbf{K} - \mathbf{K}_K. \end{aligned}$$

On $\mathbf{K}_H$ we construct a choice function C_2 using the method of proposition 5.2. On K_K we ensure that the choice function selects an eigenvector of G_τ. The eigenvectors of G_τ are contained in K_K and because they form an orthonormal set, we must have at least one eigenvector w of G_τ for which $| < f, w > | \geq \frac{1}{\sqrt{N}}$. Hence all sets $\mathbf{E}'[f]$ for $f \in \mathbf{K}_K$ must contain at least one eigenvector. We construct the equivalence classes of the relation R_3 on $\mathbf{K}_K$ defined by $f\ R_3\ h$ if and only if $\mathbf{E}[f] = \mathbf{E}[h]$. By the axiom of choice, we can select an eigenvector w of G_τ from each equivalence class. The choice function C then associates to members of each equivalence class the selected eigenvector w.

We now show the only if part. Suppose first that $\mathcal{H}$ is a complex space and let w be an eigenvector of G_τ that does not belong to the dictionary. If $f = w$ we have $\mathbf{E}[f] = \mathbf{E}[G_\tau f]$, so the only possible values of $C(\mathbf{E}[f])$ are the other eigenvectors of G_τ, all of which are perpendicular to w. Thus the only values of $C(\mathbf{E}[f])$ which preserve the commutativity do not satisfy $|< f, C(\mathbf{E}[f]) >| \geq \alpha \sup_{\gamma \in \Gamma} |< f, g_\gamma >|$ for any $\alpha > 0$. Hence $\mathcal{D}$ must contain all eigenvalues w of G_τ.

□

We cannot extend this result to groups generated by two non-commuting elements, such as the set of all unit translations and modulations, because non-commuting operators have different eigenvectors. For general groups when $\mathcal{H}$ has a finite dimension and $\mathcal{D}$ is a finite dictionary, we set the optimality factor $\alpha = 1$. Then $f \in \mathbf{K}$ if and only if there exists $g_\gamma \in \mathcal{D}$ and G_τ such that

$$\lambda(f) = | < f, g_\gamma > | = | < f, G_{\tau^{-1}} g_\gamma > |.$$

This set $\mathbf{K}$ is of measure 0 in $\mathcal{H}$. If for all $n \in \mathbf{N}$ $R^n f$ is not in $\mathbf{K}$, the proof of proposition 5.1 proves that the commutativity relation (5.2) remains valid for f. If the set of such functions is of measure 0 in $\mathcal{H}$ we say that the matching pursuit commutes with operators in $\mathcal{G}$ almost everywhere in $\mathcal{H}$.

Chapter 6

Chaos in Matching Pursuits

Each iteration of a matching pursuit is a solution of an M-optimal approximation problem where $M = 1$. Hence the pursuit exhibits some of the same instabilities in its choice of dictionary vectors as solutions to the M-optimal approximation problem. In this chapter we study these instabilities and prove that for a particular dictionary the pursuit is chaotic.

6.1 Renormalized Matching Pursuits

We renormalize the residues $R^n f$ to prevent the convergence of residues to zero so we can study their asymptotic properties. Let $R^n f$ be the residue after step n of a matching pursuit. The renormalized residue $\tilde{R}^n f$ is

$$\tilde{R}^n f = \frac{R^n f}{||R^n f||}. \tag{6.1}$$

The *renormalized matching pursuit map* is defined by

$$M(\tilde{R}^n f) = \tilde{R}^{n+1} f. \tag{6.2}$$

Since $R^{n+1} f = R^n f - < R^n f, g_{\gamma_n} > g_{\gamma_n}$ and

$$||R^{n+1} f||^2 = ||R^n f||^2 - | < R^n f, g_{\gamma_n} > |^2,$$

we derive that if $| < \tilde{R}^n f, g_{\gamma_n} > | \neq 1$

$$M(\tilde{R}^n f) = \tilde{R}^{n+1} f = \frac{\tilde{R}^n f - < \tilde{R}^n f, g_{\gamma_n} > g_{\gamma_n}}{\sqrt{1 - | < \tilde{R}^n f, g_{\gamma_n} > |^2}}. \tag{6.3}$$

We set $M(\tilde{R}^n f) = 0$ if $| < \tilde{R}^n f, g_{\gamma_n} > | = 1$.

At each iteration the renormalized matching pursuit map removes the largest dictionary component of the residue and renormalizes the new residue. This action is much like that of a binary shift operator acting on a binary decimal: the shift operator removes the most significant digit of the expansion and then multiplies the decimal by 2, which is analogous to a renormalization.

Definition 6.1 *Let $s \in [0,1]$ be expanded in binary form $0.s_1 s_2 s_3 \ldots$, where $s_i \in \{0,1\}$. The binary left-shift map $L : [0,1] \rightarrow [0,1]$ is defined by*

$$L(0.s_1 s_2 s_3 \ldots) = 0.s_2 s_3 s_4 \ldots. \tag{6.4}$$

The binary shift map is well-known to be chaotic with respect to the Lebesgue measure on $[0,1]$. We recall the three conditions that characterize a chaotic map $T : \Sigma \rightarrow \Sigma$ [14] [10].

1. T must have a *sensitive dependence on initial conditions.* Let $T^{(k)} = T \circ T \circ \ldots \circ T$, k times. There exists $\epsilon > 0$ such that in every neighborhood of $x \in \Sigma$ we can find a point y such that $|T^{(k)}(x) - T^{(k)}(y)| > \epsilon$ for some $k \geq 0$.

2. Successive iterations of T must mix the domain. T is said to be *topologically transitive* if for every pair of open sets $U, V \subseteq \Sigma$, there is a $k > 0$ for which $T^{(k)}(U) \cap V \neq \emptyset$.

3. The *periodic points* of T must be *dense* in Σ.

The topological properties of the renormalized matching pursuit map are similar to those of the left shift map which suggests the possibility of chaotic behavior. The renormalized matching pursuit map has "sensitive dependence" on the initial signal f, when f is near a dictionary element or at the midpoint of a line joining two different dictionary elements. Let $f \in \mathcal{H}$ and g_{γ_1} and g_{γ_2} be two dictionary elements such that

$$| < f, g_{\gamma_1} > | = | < f, g_{\gamma_2} > | > | < f, g_\gamma > | \quad \text{for } \gamma_1, \gamma_2 \neq \gamma \in \mathbf{\Gamma}.$$

We can change the residue Rf completely by moving f an arbitrarily small distance towards either g_{γ_1} or g_{γ_2}. The map thus separates points in particular regions of the space. Alternatively, consider two signals f_1 and f_2 defined by

$$f_1 = (1-\epsilon) g_{\gamma_0} + \epsilon h_1 \tag{6.5}$$

and

$$f_2 = (1-\epsilon) g_{\gamma_0} + \epsilon h_2 \tag{6.6}$$

where g_{γ_0} is the closest dictionary element to f_1 and f_2, $\|h_1 - h_2\| = 1$, and $< h_1, g_{\gamma_0} >=< h_2, g_{\gamma_0} >= 0$. Then $\|f_1 - f_2\| = \epsilon \|h_1 - h_2\|$ can be made arbitrarily small, while $\|\tilde{R} f_1 - \tilde{R} f_2\| = \|h_1 - h_2\| = 1$. The open ball around g_{γ_0} is mapped to the entire orthogonal complement of g_{γ_0} in the function space, which shows that in some regions of the space, the renormalized matching pursuit map also shares the domain-mixing properties of chaotic maps.

6.2 Chaotic Three-Dimensional Matching Pursuit

To prove analytically that a non-linear map is topologically transitive is often extremely difficult. We thus concentrate first on a simple dictionary of $\mathcal{H} = \mathbf{R}^3$, where we prove that the renormalized matching pursuit is topologically equivalent to a shift map. The dictionary $\mathcal{D}$ consists of three unit vectors g_0, g_1, and g_2 in $\mathbf{R}^3$ oriented such that $< g_i, g_j >= \frac{1}{2}$ for $i \neq j$. The vectors form the edges of a regular tetrahedron emerging from a common vertex; each vector is separated by a 60 degree angle from the other two.

To prove the topological equivalence, we first reduce the normalized matching pursuit to a one-dimensional map. The residue $R^n f$ is formed by projecting $R^{n-1} f$ onto the plane perpendicular to the selected $g_{\gamma_{n-1}}$. Hence, the residues $R^n f$ are all contained in one of the three planes P_i orthogonal to the vectors g_i. We can expand the residue $R^n f \in P_i$ over the orthonormal basis $(e_{i,1}, e_{i,2})$ of P_i given by

$$\begin{aligned} e_{i,1} &= g_{i+1} - g_{i-1} && (6.7) \\ e_{i,2} &= \frac{g_{i+1} + g_{i-1} - g_i}{\sqrt{2}}. && (6.8) \end{aligned}$$

All subscripts above and for the remainder of this section will be taken modulo 3.

Let (x_n, y_n) be the coordinates of $R^n f$ in the basis $(e_{i,1}, e_{i,2})$. Since it is orthogonal to g_i the next dictionary vector that is selected is either g_{i-1} or g_{i+1}. One can verify that the residue $R^n f$ is mapped to a point in P_{i-1} if $x_n y_n \leq 0$ and to a point in P_{i+1} if $x_n y_n \geq 0$. The coordinates of the residue $R^{n+1} f$ is either is these planes are

$$F_{xy} \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{cases} \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & 0 \end{bmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} & \begin{matrix} x_n > 0, y_n \geq 0 \\ \text{or } x_n < 0, y_n \leq 0 \end{matrix} \\ \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 0 \end{bmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} & \begin{matrix} x_n \geq 0, y_n < 0 \\ \text{or } x_n \leq 0, y_n > 0 \end{matrix} \end{cases} \qquad (6.9)$$

The normalized residue $\tilde{R}^n f$ has a unit norm and hence lies on a unit circle in one of the planes P_i. We can thus parameterize this residue by an angle $\theta \in [-\pi, \pi)$ with respect to the orthogonal basis $(e_{i,1}, e_{i,2})$. The angle of the next renormalized residue $\tilde{R}^{n+1} f$ in P_{i+1} or P_{i-1} is $F(\theta) = \mathrm{Arg}(F_{xy}(\cos\theta, \sin\theta))$. The graph of $F(\theta)$ is shown in Figure 6.1. To simplify the analysis, we identify the three unit circles on the planes P_i to a single circle so that the map $F(\theta)$ becomes a map from the unit circle onto itself. The index of the plane in which a residual vector $R^n f$ lies can be obtained from the index of the plane P_i in which Rf lies and the sequence of the angles in the planes of the residues $Rf, R^2 f, R^3 f, \ldots$, so the map encodes the plane P_i containing $R^n f$.

F is piecewise strictly monotonically increasing with discontinuities at integer multiples

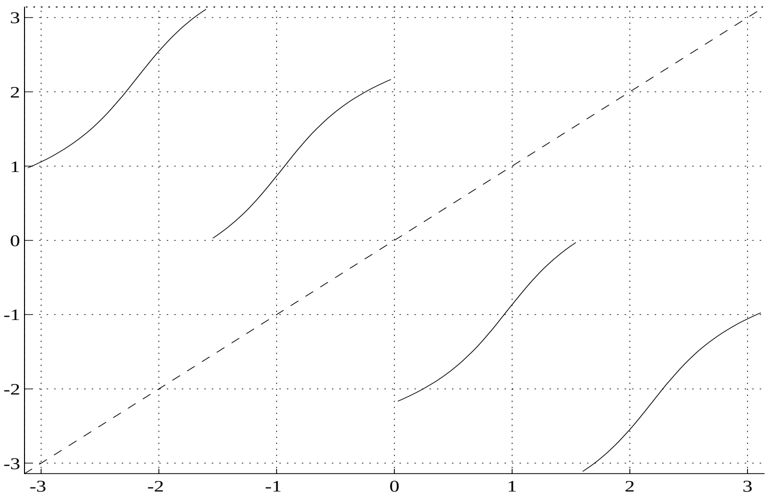

Figure 6.1: $F(\theta)$ on $[-\pi, \pi)$. The discontinuities occur between quadrants and correspond to the points at which the element selected by the pursuit changes. The first and third pieces are mapped to P_{i+1} and the second and fourth are mapped to P_{i-1}. The line $y = \theta$ is plotted for reference.

of $\frac{\pi}{2}$. Moreover, F possesses the following symmetries which we will use later:

$$F(\theta) = \begin{cases} \pi + F(\theta + \pi) & -\pi \leq \theta < -\frac{\pi}{2} \\ -F(-\theta) & -\frac{\pi}{2} \leq \theta < 0 \\ F(\theta) & 0 \leq \theta < \frac{\pi}{2} \\ \pi - F(\pi - \theta) & \frac{\pi}{2} \leq \theta < \pi \end{cases} \tag{6.10}$$

To analyze the chaotic behavior of this map, we focus on $F^{(2)}$ which is shown in Figure 6.3. The map $F^{(2)}$ partitions $[-\pi, \pi)$ into four invariant sets $I_+ = [p_1, p_2)$, $I_- = [-p_2, -p_1)$, $J_+ = [0, p_1) \cup [p_2, \pi)$, and $J_- = [-\pi, -p_2) \cup [-p_1, 0)$. Here $\pm p_1$ and $\pm p_2$ are the four fixed points of the map given by $p_1 = \tan^{-1}(\sqrt{2})$ and $p_2 = \pi - \tan^{-1}(\sqrt{2})$.

Proposition 6.1 *The restriction of $F^{(2)}$ to each of the invariant regions $I\pm$ and $J\pm$ is topologically conjugate to the binary shift map L and are therefore chaotic. Hence F is chaotic on the inverse images of $I\pm$ and $J\pm$.*

Proof: To prove that $F^{(2)}$ is *topologically conjugate* to L, we must construct an homeomorphism h such that

$$h \circ F^{(2)} = L \circ h. \tag{6.11}$$

This homeomorphism guarantees that $F^{(2)}$ shares the shift map's topological transitivity, sensitive dependence on initial conditions, and dense periodic points.

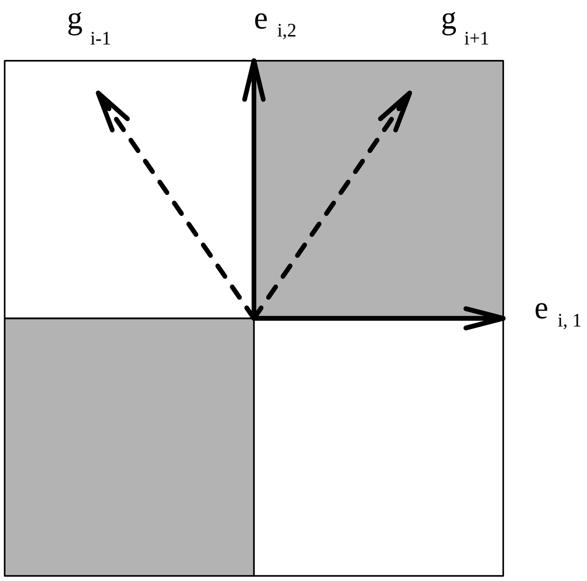

Figure 6.2: The plane P_i. The projections of g_{i+1} and g_{i-1} are shown in the first and second quadrants. The darkly and lightly shaded areas are mapped to P_{i+1} and P_{i-1}, respectively, by the next iteration of the pursuit.

We first focus on the region I_+. Due to the symmetry, the construction is identical for I_-, and we drop the subscript below. The map $F^{(2)}$ is differentiable over $I_0 = [p_1, \frac{\pi}{2}[$ and $I_1 = [\frac{\pi}{2}, p_2)$. For x in I, we define the index of x by

$$i(x) = \begin{cases} 0, & x \in I_0 \\ 1, & x \in I_1 \end{cases} \tag{6.12}$$

The *itinerary* of a point $x \in I$ is the sequence of indices of the images of x under successive applications of $F^{(2)}$. Following a standard technique [14], the homeomorphism h is constructed by assigning to each point $x \in I$ a binary decimal in $[0,1]$ with digits corresponding to the itinerary of x

$$h(x) = 0\ .\ i(x)\ i(F^{(2)}(x))\ i(F^{(4)}(x))\ i(F^{(6)}(x)) \ldots \tag{6.13}$$

The itinerary of $F^{(2)}(x)$ is just the itinerary of x shifted left, so we have

$$\begin{aligned} h \circ F^{(2)}(x) &= 0\ .\ i(F^{(2)}(x))\ i(F^{(4)}(x))\ i(F^{(6)}(x)) \ldots \\ &= L \circ h(x). \end{aligned} \tag{6.14}$$

Thus, (6.11) is satisfied. The details of the proof that $h(x)$ is a homeomorphism are similar to [14], §1.7, with one minor difference. The fact that h is one-to-one in [14] requires that $(F^{(2)})'$ be bounded above one. This is not the case here. However, we can show that

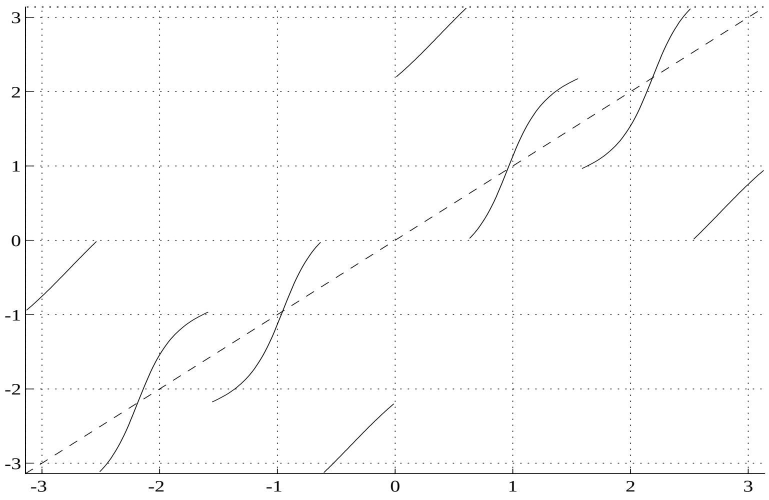

Figure 6.3: $F^{(2)}(\theta)$ on $[-\pi, \pi)$. The discontinuities again correspond to the different selected elements in the two iterations of the pursuit. From left to right in $[-\pi, 0)$, the pieces correspond to selecting (1) g_{i+1} followed by g_{i+2}, (2) g_{i+1} followed by g_i, (3) g_{i-1} followed by g_i, (4) g_{i-1} followed by g_{i-2}, with the cycle repeated in $[0, \pi)$. The fixed points correspond to the projections of $\pm g_{i\pm1}$ onto P_i.

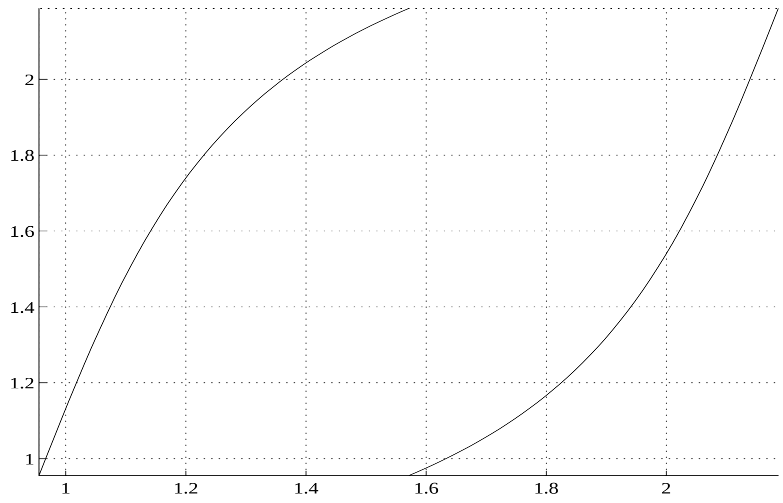

Figure 6.4: $F^{(2)}(\theta)$ on I.

for $\theta \in [-\pi, \pi)$ $(F^{(4)})'(\theta) \geq \frac{11}{6} > 1$. The injectivity of h is then obtained with minor modifications of the original proof.

The proof that $F^{(2)} : J_\pm \to J_\pm$ is a chaotic map is similar to the proof for $F^{(2)} : I_\pm \to I_\pm$. We consider J_+. We first modify the metric over our domain so that the points p_1 and p_2 have a zero distance as well as the points 0 and π. This metric over our domain is equivalent to a uniform metric over a circle. With this modification, we obtain a map which is differentiable over $[0, p_1)$ and $[p_2, \pi)$ and which maps each of these intervals to the entire domain. The proof now proceeds exactly as above. We note that in the proof that $F^{(2)}$ is chaotic on J we define the index function $i(x)$ so that

$$i(x) = \begin{cases} 0, & x \in F(I_0) \\ 1, & x \in F(I_1) \end{cases} \tag{6.15}$$

With this construction we obtain a conjugacy between F and the shift map with a homeomorphism $h'(x) = h(F(x))$.

□

The similarities of $F^{(2)}$ and L become much clearer when we compare the graph of $F^{(2)}$ on I in Figure 6.5 with the graph of the binary shift L on $[0, 1)$, given by $y = 2x \bmod 1$. Both maps are piecewise differentiable and monotonically increasing, and both map each continuous piece onto the entire domain. The slope of the graph of L is strictly greater than 1, and although the slope of the pieces of $F^{(2)}$ is not everywhere greater than 1, the slope of the pieces of $F^{(4)}$ is. The itinerary for a point in $[0, 1)$ under L is just its binary

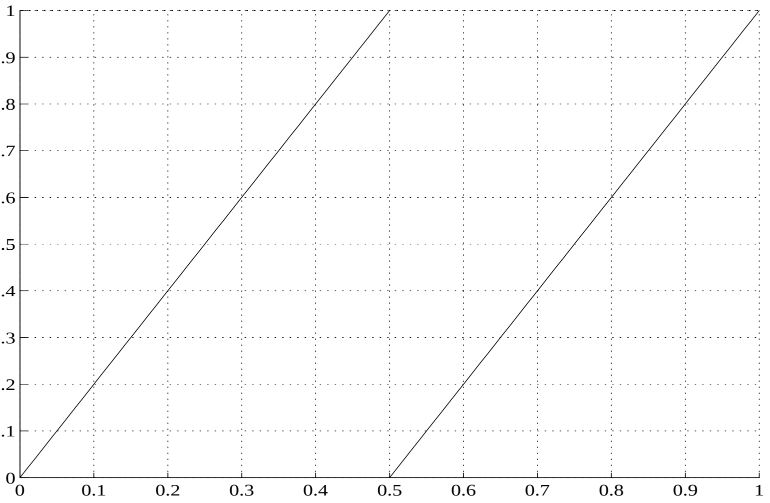

Figure 6.5: The binary left shift operator L on binary expansions of $[0,1]$.

decimal expansion, so we see that the homeomorphism we have constructed is a natural one.

Chapter 7

Invariant Measure

The chaotic properties of matching pursuits make it impossible to predict the exact evolution of the residues, but we can can obtain a statistical description of the properties of the residues. For an ergodic map, asymptotic statistics can be obtained from the invariant measure. The residue can then be interpreted as a realization of an equilibrium process whose distribution is characterized by the invariant measure of the map. The next section describes the basic properties of these invariant measures and analyzes the particular case of the three-dimensional dictionary.

In higher dimensional spaces, numerical experiments show that the norm of the residues $||R^n f||$ decreases quickly for the first few iterations, but afterwards the decay rate slows down and remains approximately constant. The average decay rate can be computed from the invariant measure and the measurement of this decay rate has applications for the approximation of signals using a small number of "coherent structures".

Families such as the Gabor dictionary that are invariant under the action of group operators yield invariant measures with invariant properties as described in chapter 5. To refine our understanding of the invariant measures, we construct an approximate stochastic model of the equilibrium process and provide numerical verifications for a dictionary composed of discrete Dirac and Fourier bases.

7.1 Ergodicity

We first summarize some results of ergodic theory [28] [47]. Let μ be a measure and let Σ be a measurable set with $\mu(\Sigma) > 0$. Let T be a map from Σ onto Σ. T is said to be *measure-preserving* if for any measurable set $S \subset \Sigma$ we have

$$\mu(S) = \mu(T^{-1}(S)), \tag{7.1}$$

where $T^{-1}(S)$ is the inverse image of S under T. The measure μ is said to be an *invariant measure* under T. A set E is said to be an *invariant set* under T if $T^{-1}E = E$. The measure preserving map T is *ergodic* with respect to a measure μ if for all invariant sets $E \subset \Sigma$ we have either $\mu(E) = 0$ or $\mu(\Sigma - E) = 0$.

Ergodicity is a measure-theoretical notion that is related to the topological transitivity property of chaos [56]. It implies that the map T mixes around the points in its domain. For example, if T has an ergodic invariant measure μ that is non-atomic (every set of non-zero measure contains a subset of smaller measure), then only for a set of μ-measure 0 do the iterates $Tx, T^2x, T^3x, \ldots$ converge to a cycle of finite length. Hence, for almost all $x \in \Sigma$, $T^n x$ neither goes to a fixed point or a limit cycle, so for most of Σ the asymptotic behavior of $T^n x$ is complicated.

The binary left shift map on [0,1] is ergodic with respect the Lebesgue measure [38]. We can use the topological conjugacy relation (6.11) we derived in chapter 6 to prove that the renormalized matching pursuit map F is also ergodic when restricted to one of two invariant sets. We first prove that the map $F^{(2)}$ of the previous chapter is ergodic.

Lemma 7.1 *The restrictions of $F^{(2)}$ to the invariant sets $I_\pm, J_\pm$ are ergodic.*

Proof: Let Σ be one of the sets $I_\pm, J_\pm$. We first show that $F^{(2)}$ is measure-preserving. Let h be the homeomorphism satisfying

$$h \circ F^{(2)} = L \circ h \tag{7.2}$$

on Σ, and let ν be the Lebesgue measure. The shift map L preserves the Lebesgue measure, so we have for any set $S \in [0, 1]$ that

$$\nu(S) = \nu(L^{-1}S). \tag{7.3}$$

We define $\mu(S) = \nu(h(S))$, where $h(S) = \{h(x) : x \in S\}$. This μ is a measure because h is a measurable function. We have

$$\begin{aligned}
\mu(S) &= \nu(h(S)) \\
&= \nu(L^{-1} \circ h(S)) \\
&= \nu(h \circ (F^{(2)})^{-1} \circ h^{-1} \circ h(S)) \\
&= \mu((F^{(2)})^{-1}(S)), \qquad (7.4)
\end{aligned}$$

so $F^{(2)}$ preserves the measure μ.

Suppose that the set S is invariant with respect to $F^{(2)}$, i.e. $F^{(2)}(S) = S$. From (7.3), the set $h(S)$ must be invariant under L, and because L is ergodic, either the Lebesgue measure of $h(S)$ or the Lebesgue measure of $([0, 1] - h(S))$ must be zero. By our construction of ν, though, we must then have that $\mu(S) = 0$ or $\mu(\Sigma - S) = 0$. Hence $F^{(2)}$ is ergodic.

□

Proposition 7.1 *The three-dimensional renormalized matching pursuit map F is ergodic when restricted to one of its two invariant sets.*

Proof: The map F is invariant on the sets $I_- \cup J_+$ and $J_- \cup I_+$, and we have $F(I_\pm) = J_\mp$ and $F(J_\pm) = I_\mp$. Let Σ be one of the two invariant sets. To simplify our notation, we will drop the the subscripts of I and J.

We first show that F is measure-preserving. We define μ_I and μ_J to be the ergodic invariant measures on I and J, respectively, which were derived in the above lemma. We have $\mu_I(I) = \mu_J(J) = 1$ from our derivation in the lemma. We can decompose any set $S \subset \Sigma$ into the disjoint union of $S_I = S \cap I$ and $S_J = S \cap J$. Let χ_S be the characteristic function of S. By the Birkhoff ergodic theorem,

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{2n} \sum_{k=0}^{2n-1} \chi_S(F^{(k)}(x)) &= \lim_{n\to\infty} \frac{1}{2n} \sum_{k=0}^{n-1} \chi_S(F^{(2k)}(x)) + \chi_S(F^{(2k+1)}x) \\ &= \frac{1}{2}[\mu_I(S_I) + \mu_I(F^{-1}(S_J)) + \mu_J(S_J) + \mu_J(F^{-1}(S_I))] \\ &= \mu(S) \end{aligned} \tag{7.5}$$

for almost all x. μ is a measure which is invariant with respect to F due to the invariances of μ_I and μ_J under $F^{(2)}$, and $\mu(\Sigma) = 1$.

We now show that μ is ergodic. Let S is a μ measurable set which is invariant under F, and let Σ' be the subset of Σ of full measure for which the sum in (7.5) converges. We have $\mu(S) = \mu(S \cap \Sigma')$. Suppose that $\mu(S \cap \Sigma') > 0$. Then we must have

$$\mu(S \cap \Sigma') = \lim_{n\to\infty} \frac{1}{2n} \sum_{k=0}^{2n-1} \chi_{S\cap\Sigma'}(F^{(k)}(x)) = 1 = \mu(\Sigma). \tag{7.6}$$

Either $\mu(S) = 0$ or $\mu(\Sigma - S) = 0$, so the result is proved.

□

The ergodicity of a map T allows us to numerically estimate the invariant measure by counting for points $x \in \Sigma$ how often the iterates $Tx, T^2x, T^3x, \ldots$ lie in a particular subset S of Σ. The Birkhoff ergodic theorem [28] states that when $\mu(\Sigma) < \infty$,

$$\mu(S) = \mu(\Sigma) \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \chi_S(T^k x). \tag{7.7}$$

When an invariant measure μ is absolutely continuous with respect to the Lebesgue measure, by the Radon-Nikodym theorem there exists a function p such that

$$\mu(S) = \int_S p(x) \tag{7.8}$$

The function p is called an *invariant density*. For the invariant measure of F, this density is given by

$$p(x) = |h'(x)|$$

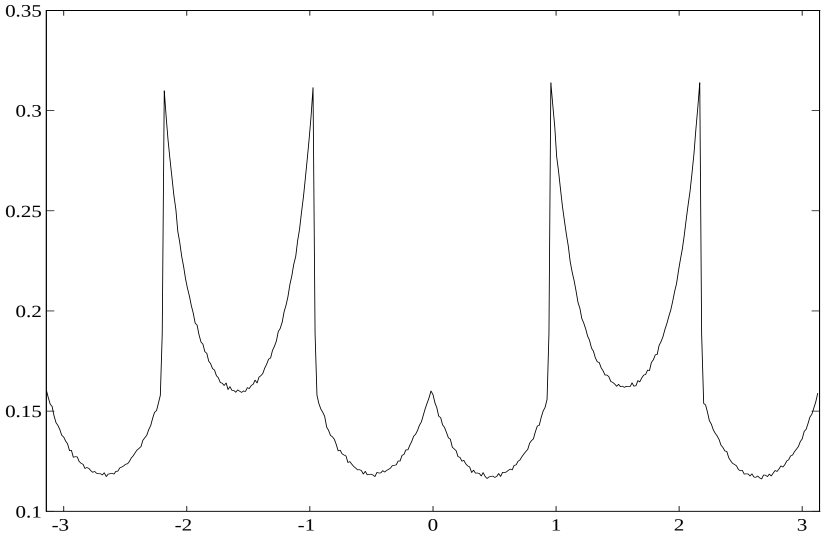

Figure 7.1: The invariant densities of F for the two invariant sets $I_{\pm} \cup J_{\mp}$ superimposed on the interval $[-\pi, \pi)$. The densities have been obtained by computing the Cesaro sums.

provided that $h(x)$ is absolutely continuous. This invariant density measure can be computed numerically by estimating the limit (7.7) when the density exists. Fig. 7.1 is the result of numerically computing the Cesaro sums (7.7) for a large set of random values of x with sets S of the form $[a, a + \delta_a)$. In this case, the support of the invariant measure of the normalized matching pursuit is on the three unit circles of the planes P_i. On each of these circles, the invariant density measures are the same and equal to $p(\theta)$.

7.2 Coherent Structures

The Cesaro sum (7.7) shows that an ergodic invariant measure reflects the distributions of iterates of the map T. The average number of times the map takes its value in a set is proportional to the measure of this set. The invariant measure thus provides a statistical description after a large number of iterations, during which the map may have transient behavior. For example, for the three-dimensional dictionary of chapter 6, there is one chance in three that the residue is on the unit circle of any particular plane P_i, and over this plane the probability that it is located at the angle θ is $p(\theta)$.

In higher dimensional spaces the invariant measure μ can be viewed as the distribution of a stochastic process over the unit sphere $\mathcal{S}$ of the space $\mathcal{H}$. After a sufficient number of iterations, the residue of the map can be considered as a realization of this process. We call "dictionary noise" the process P corresponding to the invariant ergodic measure of the renormalized matching pursuit (if it exists). If the dictionary is invariant under translations and frequency modulations, we prove in the next section that the dictionary noise is a stationary white noise. Realizations of a dictionary noise have inner products that are as small and as uniformly spread across the dictionary vectors as possible. Indeed, the measure is invariant under the action of the normalized matching pursuit which sets to zero the largest inner product and makes the appropriate renormalization with (7.16). Since the statistical properties of a realization x of P are not modified by setting to zero the largest inner product, the value $\lambda(x)$ of this largest inner product cannot be much larger the next larger ones. The average value of this maximum inner product for realizations of this process is by definition

$$\lambda_\infty = \int_{\mathcal{S}} \lambda(x) d\mu(x) = E[\lambda(P)].$$

The ergodicity of the invariant measure implies that

$$\lambda_\infty = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \lambda(\mathbf{R}^k f x). \qquad (7.9)$$

We recall from (3.12) that if the optimality factor $\alpha = 1$ we have

$$\|R^{n+1} f\| = \|R^n f\| \sqrt{1 - \lambda^2(R^n f)}. \qquad (7.10)$$

The average decay rate is thus

$$d_\infty = \lim_{n\to\infty} \frac{\log\|f\| - \log\|R^{n-1} f\|}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \frac{-1}{2} \log(1 - \lambda^2(\tilde{R}^k f)). \qquad (7.11)$$

The ergodicity of the renormalized map implies that this average decay rate is

$$d_\infty = -\frac{1}{2} \int_{\mathcal{S}} \log(1 - \lambda^2(x)) d\mu(x) = -\frac{1}{2} E[\log(1 - \lambda^2(P))]. \qquad (7.12)$$

Since $\lambda(x) \geq \lambda_{min}$,

$$d_\infty \geq -\frac{1}{2} \log(1 - \lambda^2_{min}),$$

but numerical experiments show that there is often not a large factor between these two values.

The decay rate of the norms of the residues $R^n f$ was studied numerically in [58]. The numerical experiments show that when the original vector f correlates well with a few dictionary vectors, the first iterations of the matching pursuit remove these highly correlated components, called *coherent structures*. Afterwards, the average decay rate decreases quickly to d_∞.

The chaotic behavior of the matching pursuit map that we have derived provides a theoretical explanation for this behavior of the decay rate. As the coherent structures are removed, the energy of the residue becomes spread out over many dictionary vectors, as it is for realizations of the dictionary noise P, and the decay rate of the residue becomes small and on average equal to d_∞. The convergence of the average decay rate to d_∞ can be interpreted as the the residues of an ergodic map converging to the support of the invariant measure.

We emphasize that our notion of coherence here is entirely dependent upon the dictionary in question. A residue which is considered dictionary noise with respect to one dictionary may contain many coherent structures with respect to another dictionary. For example, a sinusoidal wave has no coherent components in a dictionary composed of Diracs but is clearly very coherent in a dictionary of complex exponentials.

For many signal processing applications, the dictionary defines a set of structures which we wish to isolate. We truncate signal expansions after most of the coherent structures have been removed because the dictionary noise which remains does not resemble the features we are looking for, and because the convergence of the approximations is slow for the dictionary noise. Expansions into coherent structures allow us to compress much of the signal energy into a few elements.

As long as a signal f contains coherent structures, the sequence $\lambda(R^n f)$ has different properties than realizations of the random variable $\lambda(P)$, where P is the dictionary noise process. A simple procedure to decide when the coherent structures have mostly disappeared by iteration n is to test whether a running average of the $\lambda(R^k f)$'s satisfy

$$\frac{1}{d}\sum_{k=n}^{n+d} \lambda(R^k f) \leq \lambda_\infty(1+\epsilon), \tag{7.13}$$

where d is a smoothing parameter and ϵ is a confidence parameter that are adjusted depending upon the variance of $\lambda(P)$.

Numerical experiments indicate that the normalized matching pursuit with a Gabor dictionary does have an ergodic invariant measure. After a number of iterations, the residues behave like realizations of a stationary white noise. The next section shows why this occurs. In our discrete implementation of this dictionary, where the scale is discretized in powers of 2 and $\mathcal{H} = \mathbf{R}^N$ where $N = 8192$, we measured numerically that $\lambda_\infty \approx 0.043 = \frac{3.9}{\sqrt{8192}}$. Fig. 7.3 displays $\lambda(R^n f)$ as a function of the number of iterations n for a noisy recording of the word "wavelets" shown in Fig. 7.2. We see that the Cesaro average of $\lambda(R^n f)$ is converging to λ_∞. The time-frequency energy distribution $Ef(t,\omega)$ of the first $n = 200$ coherent structures is shown in Fig. 7.6. Fig. 7.5 is the signal reconstructed from these coherent structures whereas Fig. 7.7 shows the approximation error $R^n f$. The signal recovered from the coherent structures has an excellent sound quality despite the fact that it was approximated by many fewer elements than the number of samples.

When we use the Gabor dictionary, the coherent structures of a signal are those portions of a signal which are well-localized in the time-frequency plane. White noise is not efficiently represented in this dictionary because its energy is spread uniformly over the entire dictionary, much like the realizations of the dictionary noise. We analyze expansions

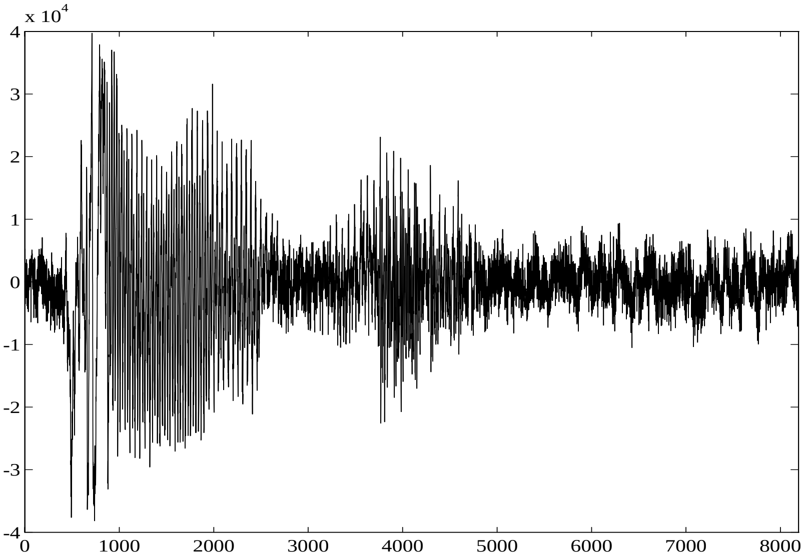

Figure 7.2: Digitized recording of a female speaker pronouncing the word “wavelets” to which white noise has been added. Sampling is at 11 KHz, and the signal to noise ratio is 10dB.

of realizations of a white noise process in detail in section 7.5. Speech contains many structures which are well-localized in the time-frequency plane, especially in voiced segments of speech, so speech signals are efficiently represented by the Gabor dictionary. The coherent portion of a noisy speech signal, therefore, will be a much better approximation to the speech than to the noise. As a result, the coherent reconstruction of the “wavelet” signal has a 14.9 dB signal to noise ratio whereas the original signal had only a 10.0 dB SNR. Moreover, the coherent reconstruction is audibly less noisy than the original signal.

[58] proposes a denoising procedure based upon the fact that white noise is poorly represented in the Gabor dictionary, which was inspired by numerical experiments with the decay of the residues. Similar ideas exist in [51] [18], namely, to separate “noise” from a signal, we approximate a signal using a scheme which efficiently approximates the portion of interest but inefficiently approximates the noise. In order to implement a denoising scheme with a matching pursuit, it is essential that the dictionary be well-adapted to decomposing that portion of signals we wish to retain and poorly-adapted to decomposing that portion we wish to discard. In chapter 8 we describe an algorithm for optimizing a dictionary so that we can maximize the coherence of signals of interest. Furthermore, the analysis of this chapter can be used to characterize the types of signals that a given dictionary is inefficient for representing, the realizations of a dictionary noise, so that we can determine what types of “noise” we can remove from signals.

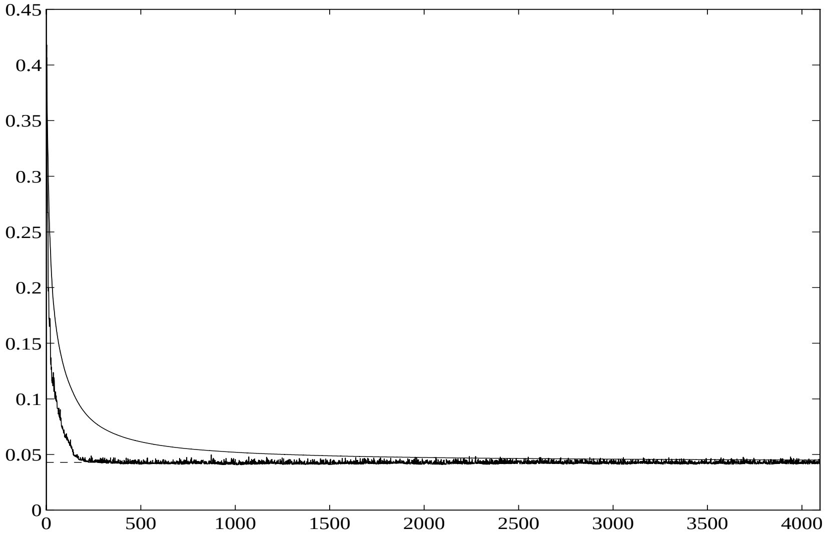

Figure 7.3: $\lambda(R^n f)$ and the Cesaro sum $\frac{1}{n}\sum_{k=1}^{n}\lambda(R^k f)$ as a function of n for the "wavelets" signal with a dictionary of discrete Gabor functions. The top curve is the Cesaro sum, the middle curve is $\lambda(R^n f)$, and the dashed line is λ_∞. We see that both $\lambda(R^n f)$ and the Cesaro sum converge to λ_∞ as n increases.

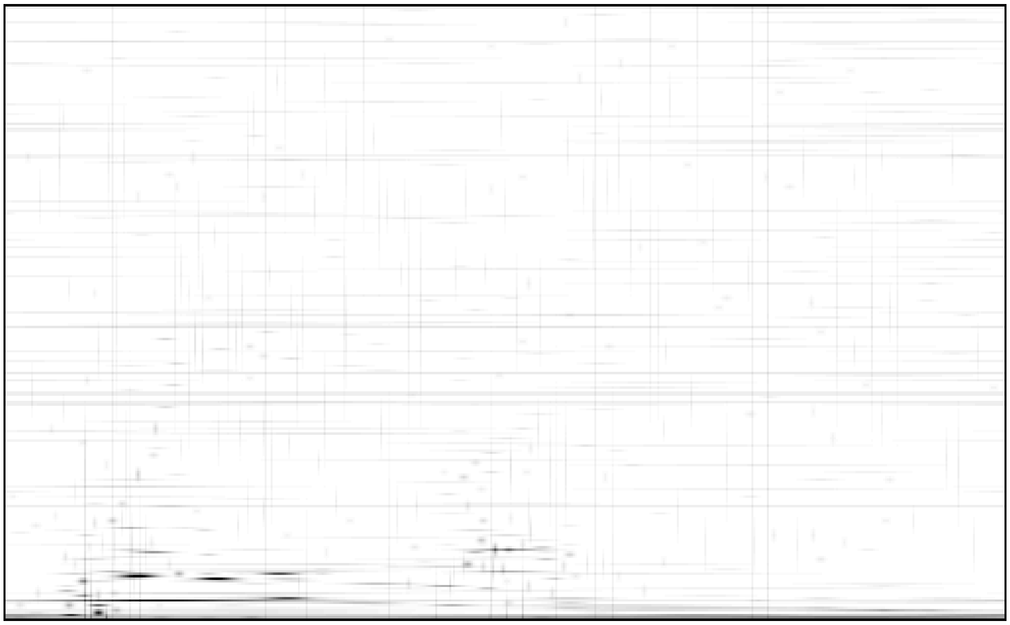

Figure 7.4: The time-frequency energy distribution of the speech recording shown in Fig. 7.2. The initial cluster which contains the low-frequency "w" and the harmonics of the long "a". The second cluster is the "le". The final portion of the signal is the "s", which resembles a band-limited noise. The faint horizontal and vertical bars scattered across the time-frequency plane are components of the Gaussian white noise which was added to the speech signal.

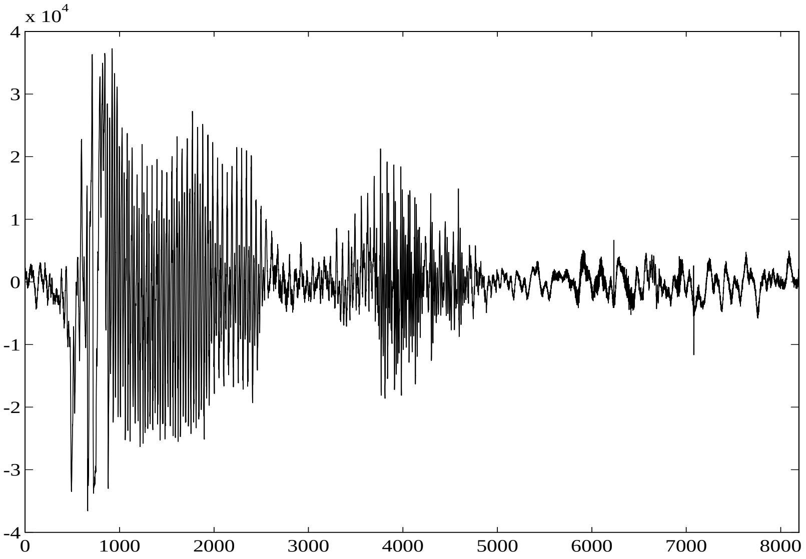

Figure 7.5: The "wavelets" signal reconstructed from the 200 coherent structures. The number of coherent structures was determined by setting $d = 5$ and $\epsilon = 0.02$.

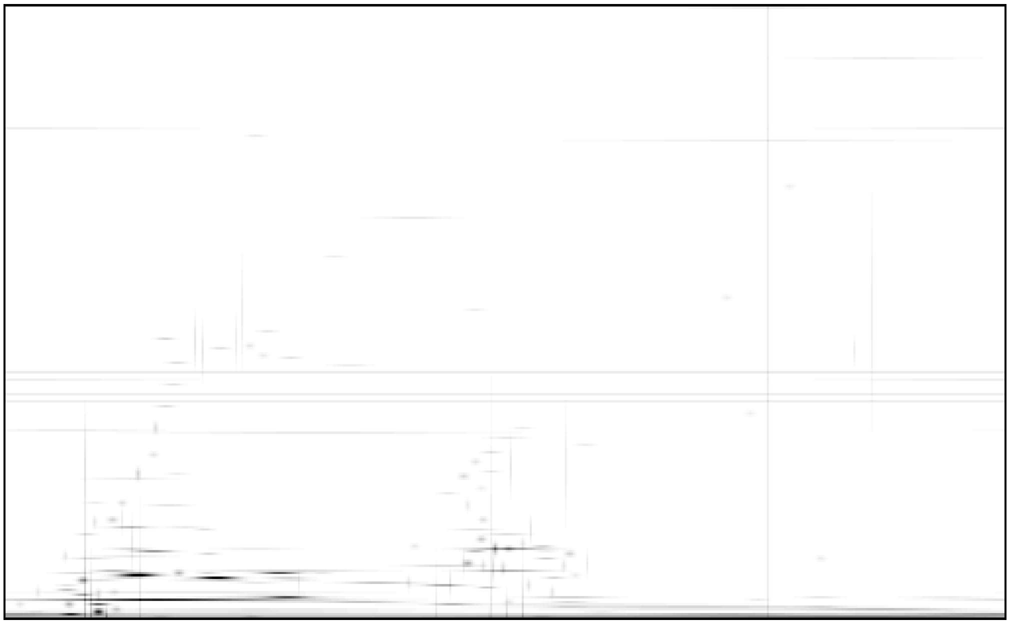

Figure 7.6: The time-frequency energy distribution of the 200 coherent structures of the speech recording shown in Fig. 7.2.

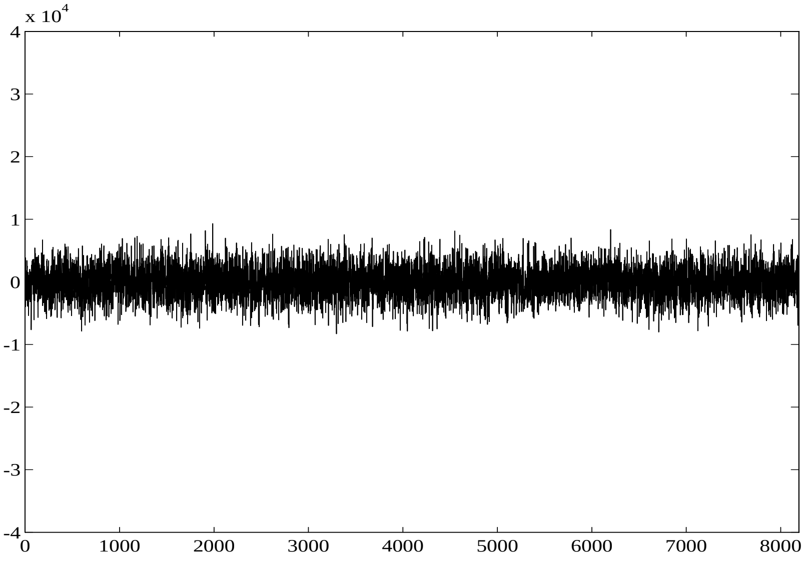

Figure 7.7: The residual $R^{200}f$ of the "wavelets" signal shown in Fig. 7.2.

7.3 Invariant Measure of Group Invariant Dictionaries

The Gabor dictionary is a particular example of a dictionary that is invariant under the action of group operators $\mathcal{G} = \{G_\tau\}_{\tau\in\Omega}$. We proved in chapter 5 that with an appropriate choice function the resulting matching pursuit commutes with the corresponding group operators. If the matching pursuit commutes with G_τ, the renormalized matching pursuit map also satisfies the commutativity property

$$M(G_\tau \tilde{R}^n f) = G_\tau M(\tilde{R}^n f). \tag{7.14}$$

The following proposition studies a consequence of this commutativity for the invariant measure, in a finite dimensional space.

Proposition 7.2 *Let M be an ergodic matching pursuit map with an invariant measure μ defined on the unit sphere $\mathcal{S}$ with $\mu(\mathcal{S}) < +\infty$. If there exists a set of $f \in \mathcal{S}$ of non-zero μ-measure such that (7.14) is satisfied for all $n \in \mathrm{N}$, then for any $G_\tau \in \mathcal{G}$ and $U \in \mathcal{S}$*

$$\mu(G_\tau U) = \mu(U).$$

Proof: This result is a simple consequence of the Birkhoff ergodic theorem. Indeed for any $U \subset \Sigma$ and almost any $f \in \mathcal{S}$ whose residues satisfy (7.14)

$$\mu(U) = \mu(\mathcal{S}) \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \chi_U(M^k f). \tag{7.15}$$

Hence

$$\mu(G_\tau U) = \mu(\mathcal{S}) \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \chi_{G_\tau U}(M^k f).$$

Since $M^k G_{\tau^{-1}} f = G_{\tau^{-1}} M^k f$

$$\chi_{G_\tau U}(M^k f) = \chi_U(M^k G_{\tau^{-1}} f).$$

But since the limit (7.15) is independent of f for almost all f, we derive that $\mu(G_\tau U) = \mu(U)$.

□

This result trivially applies to the invariant measure of the three dimensional dictionary studied in chapter 6. Since the three vectors $\{g_1, g_2, g_3\}$ have equal angles of 60 degrees between themselves, this dictionary is invariant under the action of the rotation group composed of $\{I, G, G^2\}$ where I is the identity and G the rotation operator that maps g_i to g_{i+1} where the index i is taken modulo 3. This implies that the invariant measure of the normalized matching pursuit is invariant with respect to G. It thus admits the same invariant measure over the unit circle in each plane P_i.

A more interesting application of this result concerns dictionaries that are invariant by translation and frequency modulation groups. Let $\mathcal{H} = \mathbf{R}^N$ and $\{\delta_n\}_{0\leq n<N}$ be the canonical (or Dirac) basis. The translation group is composed of $\{T^k\}_{0\leq k<N}$ where T is the elementary translation modulo N

$$T\delta_n = \delta_{(n+1) \bmod N}.$$

The modulation group is composed of $\{F^k\}_{0\leq k<N}$ where F is the frequency modulation operator defined by

$$F\delta_n = e^{i\frac{2\pi n}{N}}\delta_n.$$

Suppose that the matching pursuit is an ergodic map which admits an invariant measure and that it is implemented with a choice function that commutes almost everywhere with the translation and frequency modulation group operators. Proposition 5.1 proves that the invariant measure of M is also invariant with respect to translations T^k and frequency modulations F^k. The invariance with respect to translations means that the discrete process associated to this measure is stationary (modulo N). The invariance with respect to frequency modulation operators F^k implies that the discrete power spectrum of this process (the discrete Fourier transform of the N point autocorrelation vector) is constant. In other words, the process is a white stationary noise.

A simple example of a translation and frequency modulation invariant dictionary is constructed by aggregating the canonical basis of N discrete Diracs and the discrete Fourier orthonormal basis

$$\mathcal{D} = \{\delta_n, e_n\}_{0\leq n<N} = \{g_\gamma\}_{\gamma\in\mathbf{\Gamma}},$$

where e_n is the discrete complex exponential

$$e_n = \sum_{k=0}^{N-1} e^{\frac{i2\pi nk}{N}} \delta_k.$$

In the next section we construct a stochastic model of the matching pursuit invariant measure obtained with this dictionary.

7.4 An Invariant Measure Model

This section describes an approximate invariant measure model that we apply to the discrete Dirac-Fourier dictionary. The model is verified numerically at the end of the section. Let g_{γ_n} be the dictionary element selected on iteration n. The normalized matching pursuit map is defined by

$$\tilde{R}^{n+1} f = \frac{\tilde{R}^n f - < \tilde{R}^n f, g_{\gamma_n} > g_{\gamma_n}}{\sqrt{1 - |< \tilde{R}^n f, g_{\gamma_n} >|^2}}. \tag{7.16}$$

To find the invariant measure we consider the matching pursuit mapping of a stochastic process P^n

$$P^{n+1} = M(P^n) = \frac{P^n - < P^n, g_{P^n} > g_{P^n}}{\sqrt{1 - |< P^n, g_{P^n} >|^2}}, \tag{7.17}$$

where g_{P^n} is a random vector that takes its values over the dictionary $\mathcal{D}$ and satisfies

$$|< P^n, g_{P^n} >| = \sup_{\gamma \in \mathbf{\Gamma}} |< P^n, g_\gamma >|. \tag{7.18}$$

The invariant measure of the map corresponds to an equilibrium state in which P^{n+1} has the same distribution as P^n. For any $\gamma \in \mathbf{\Gamma}$,

$$< P^{n+1}, g_\gamma > = \frac{< P^n, g_\gamma >}{\sqrt{1 - |< P, g_{P^n} >|^2}} - \frac{< P^n, g_{P^n} >< g_{P^n}, g_\gamma >}{\sqrt{1 - |< P^n, g_{P^n} >|^2}}. \tag{7.19}$$

We recall that

$$\lambda(P^n) = |< P^n, g_{P^n} >|. \tag{7.20}$$

We suppose that in equilibrium the random variable $\lambda(P^n)$ is constant and equal to its mean, λ_∞. This is equivalent to supposing that the standard deviation of $\lambda(P)$ is small with respect to the mean, which is indeed verified numerically with several large dimensional dictionaries.

The behavior of $< P^n, g_{P^n} >< g_{P^n}, g_\gamma >$ can be divided into three cases. If $g_{P^n} = g_\gamma$, then $< P^{n+1}, g_\gamma >= 0$. If $< g_\gamma, g_{P^n} >= 0$ then (7.19) reduces to

$$< P^{n+1}, g_\gamma > = \frac{< P^n, g_\gamma >}{\sqrt{1 - \lambda_\infty^2}}. \tag{7.21}$$

Otherwise, we decompose

$$g_{P^n} = < g_{P^n}, P^n > P^n + < g_{P^n}, Q^n > Q^n.$$

Since P^n is a process whose realizations are on the unit sphere of $\mathcal{H}$, this is equivalent to an orthogonal projection onto a unit norm vector P^n plus the projection Q^n onto the orthogonal complement of P^n. We thus obtain

$$< g_{P^n}, g_\gamma > = < g_{P^n}, P^n >< P^n, g_\gamma > + < g_{P^n}, Q^n >< Q^n, g_\gamma > . \tag{7.22}$$

Inserting this equation into (7.19) yields

$$< P^{n+1}, g_\gamma > \, = \, < P^n, g_\gamma > \sqrt{1 - |< g_{P^n}, P^n >|^2} + A_\gamma^n, \tag{7.23}$$

with

$$A_\gamma^n = - \frac{< P^n, g_{P^n} >< g_{P^n}, Q^n >< Q^n, g_\gamma >}{\sqrt{1 - |< P^n, g_{P^n} >|^2}}.$$

We have from (7.22) that

$$|A_\gamma^n| = \frac{\lambda_\infty |< g_{P^n}, g_\gamma > - < g_{P^n}, P^n >< P^n, g_\gamma >|}{\sqrt{1 - \lambda_\infty^2}}. \tag{7.24}$$

If ${\lambda_\infty}^2 \ll |< g_\gamma, g_{P^n} >|^2$, then because

$$|< P^n, g_\gamma >| \leq |< P^n, g_{P^n} >| \approx \lambda_\infty,$$

we have to a first approximation that

$$|A_\gamma^n| = \frac{\lambda_\infty |< g_{P^n}, g_\gamma >|}{\sqrt{1 - \lambda_\infty^2}}. \tag{7.25}$$

Equation (7.19) is then reduced to

$$< P^{n+1}, g_\gamma > \, = \, < P^n, g_\gamma > \sqrt{1 - \lambda_\infty^2} + \frac{\lambda_\infty |< g_{P^n}, g_\gamma >| e^{i\phi_\gamma^n}}{\sqrt{1 - \lambda_\infty^2}}, \tag{7.26}$$

where ϕ_γ^n is the complex phase of A_γ^n. The three possible new cases for the evolution of $< P^n, g_\gamma >$ are summarized by

$$< P^{n+1}, g_\gamma > = \begin{cases} \frac{< P^n, g_\gamma >}{\sqrt{1 - {\lambda_\infty}^2}}, & \text{if } < g_\gamma, g_{P^n} > = 0, \\ \sqrt{1 - {\lambda_\infty}^2} < P^n, g_\gamma > + \frac{\lambda_\infty |< g_\gamma, g_{P^n} >| e^{i\phi_\gamma^n}}{\sqrt{1 - {\lambda_\infty}^2}}, & \text{if } \lambda_\infty^2 \ll |< g_\gamma, g_{P^n} >| \\ 0, & \text{if } g_\gamma = g_{P^n}. \end{cases} \tag{7.27}$$

The Dirac-Fourier dictionary is an example of dictionary for which all these simplification assumptions are valid. We observe numerically that in the equilibrium state for a space of dimension N, λ_∞ is of the order of $\frac{1}{\sqrt{N}}$ whereas the standard deviation of $\lambda(P)$ is of the order of $\frac{1}{N}$, which justifies approximating $\lambda(P)$ by its mean λ_∞. Moreover, for any distinct g_γ and g_{P^n} in this dictionary, either both vectors are in the same basis (Dirac or Fourier) and

$$< g_\gamma, g_{P^n} >= 0,$$

or both vectors are in different bases and

$$\lambda_\infty^2 \ll |< g_\gamma, g_{P^n} >| = \frac{1}{\sqrt{N}}.$$

So one of the approximations of (7.27) always applies. Because of the symmetrical positions of the Dirac and the Fourier dictionary vectors, there is an equal probability that g_{P^n} belongs to the Dirac or Fourier basis. For any fixed g_γ, the first two updating equations of (7.27) thus apply with equal frequency. We derive an average updating equation which incorporates both equations for $g_{P^n} \neq g_\gamma$,

$$< P^{n+2}, g_\gamma > - < P^n, g_\gamma >= \frac{\lambda_\infty e^{i\phi_\gamma^n}}{\sqrt{N}}. \tag{7.28}$$

For n and γ fixed, $e^{i\phi_\gamma^n}$ is a complex random variable and the symmetry of the dictionary implies that its real and imaginary parts have the same distributions with a zero mean. For any fixed γ, we also suppose that at equilibrium the phase random variables ϕ_γ^n are independent as a function of n. The difference $< P^{n+2K}f, g_\gamma > - < P^n, g_\gamma >$ is thus the sum of K independent, identically distributed complex random variables of variance 1. By the central limit theorem, the distribution of $\frac{1}{2K}(< P^{n+2K}f, g_\gamma > - < P^n, g_\gamma >)$ tends to a complex Gaussian random variable of variance 1. The inner products $< P^n, g_\gamma >$ thus follow a complex random walk as long as $g_\gamma \neq g_{P^n}$. The last case $g_{P^n} = g_\gamma$ of (7.27) occurs when $< P^n, g_\gamma >$ is the largest inner product whose amplitude we know to be

$$|< P^n, g_{P^n} >| = \lambda(P) = \lambda_\infty.$$

At equilibrium, the distribution of $< P^n, g_\gamma >$ is that of a random walk with an absorbing boundary at λ_∞.

To find an explicit expression for the distribution of the resulting process, we approximate the difference equation with a continuous time Langevin differential equation

$$\frac{d}{dt} < P^t, g_\gamma >= \frac{\lambda_\infty}{2\sqrt{N}}\eta(t), \tag{7.29}$$

where $\eta(t)$ is a complex Weiner process with mean 0 and variance 1. The corresponding Fokker-Planck equation [22] describes the evolution of the probability distribution $p(z,t)$

of $z =< P^t, g_\gamma >$. Since the real and complex parts of $\eta(t)$ have same variance, the solution can be written $p(z,t) = p(r,t)$ where $r = |z|$ and

$$\frac{\partial p(r,t)}{\partial t} = \frac{{\lambda_\infty}^2}{8N} \triangle p(r,t) \tag{7.30}$$

which reduces to

$$\triangle p(r) = 0 \tag{7.31}$$

at equilibrium. The general solution to (7.31) with a singularity at $r = 0$ is

$$p(r) = C \ln(r) + D.$$

The constants C and D are obtained from boundary conditions.

The inner products $< P^n, g_\gamma >$ start at $r = 0$ and diffuse outward until they reach $r = \lambda_\infty$, at which time $g_{P^n} = g_\gamma$, and $< P^n, g_\gamma >$ returns to 0. The Langevin equation (7.29) describes the evolution of the inner products before selection; the selection process is modeled by the boundary conditions.

We can write (7.30) in the form of a local conservation equation,

$$\frac{\partial p(r,t)}{\partial t} + \frac{\partial J(r,t)}{\partial r} = 0, \tag{7.32}$$

where J, the probability current, is given by

$$J = \frac{-{\lambda_\infty}^2}{8N} \nabla p. \tag{7.33}$$

The aggregate evolution of the inner products is described by a net probability current which flows outward from a source at the origin and which is removed by a sink at $r = \lambda_\infty$. At each time step, exactly one of the $2N$ dictionary elements is selected and set to 0. Thus, the strength of both the sink and the source is $\frac{1}{2N}$ and thus implies that

$$\lim_{r \to 0} \oint_{|z|=r} J \cdot \hat{n} \; d\ell = \frac{1}{2N} \tag{7.34}$$

$$\oint_{|z|=\lambda_\infty} J \cdot \hat{n} \; d\ell = \frac{1}{2N} \tag{7.35}$$

Integrating (7.31) we find that $rp_r(r) = C$. By performing the line integrals in (7.34) and (7.35), we find that $C = \frac{-2}{\pi {\lambda_\infty}^2}$. Thus, we have

$$p(r) = \frac{-2}{\pi {\lambda_\infty}^2} \ln(r) + D. \tag{7.36}$$

We use additional constraints to find D and λ_∞. Since all inner products lie in $|r| < \lambda_\infty$, we must have

$$\int_{|z|<\lambda_\infty} p(z) dz = 1. \tag{7.37}$$

Since the dictionary includes two orthonormal bases and $\|P^t\| = 1$, we have

$$\sum_{\gamma \in \mathbf{\Gamma}} |< P^t, g_\gamma >|^2 = 2.$$

The $2N$ inner products $< P^t, g_\gamma >$ with dictionary elements correspond to particles of $2N$ different ages (where a particle's age is the time since it was last set to zero). We thus assume the mean ergodic property

$$E|< P^t, g_\gamma >|^2 = \frac{1}{2N} \sum_{\gamma \in \mathbf{\Gamma}} |< P^t, g_\gamma >|^2 = \frac{1}{N}$$

and hence

$$\int_{|z| < \lambda_\infty} z^2 p(z) dz = \frac{1}{N}. \tag{7.38}$$

Inserting conditions (7.37) and (7.38) into (7.36) yields

$$\lambda_\infty = \frac{2}{\sqrt{N}} \quad \text{and} \quad D = \frac{2 \ln \lambda_\infty}{\pi {\lambda_\infty}^2}.$$

Hence

$$p(r) = \frac{2}{\pi {\lambda_\infty}^2} \ln(\frac{\lambda_\infty}{r}). \tag{7.39}$$

Figure 7.8 compares the graph of (7.39) for $N = 4096$ with an empirically determined density function. The empirical density function was obtained by computing the Cesaro sums $\frac{1}{n} \sum_{k=0}^{n} < R^k f, g_\gamma >$ where g_γ is a Dirac element and f is a realization of a white noise. The first N terms were discarded to eliminate transient behavior and to speed the convergence of the sum. We have aggregated the Cesaro sums for the members of the Dirac basis to obtain higher resolution. The invariant density function is invariant by translation due to the translation invariance of the decomposition, so this aggregation does not affect our measurements. The figure shows an excellent agreement between the model and measured values. Figure 7.9 compares predicted values of λ_∞ with empirically determined values. The discrepancy near the origin is due to the fact that the approximation of the of the complex exponential term in (7.28) with a Gaussian is not valid for the first few iterations after $< P^n, g_\gamma >$ is set to 0. These results justify *a posteriori* the validity our approximation hypotheses.

For this dictionary the average value λ_∞ is only twice as large as the minimum λ_{min}. The value λ_{min} is attained for the linear chirp

$$f = \sum_{k=0}^{N-1} e^{\frac{i2\pi k^2}{N}} \delta_k,$$

where

$$\lambda_{min} = \lambda(f) = \frac{1}{\sqrt{N}}.$$

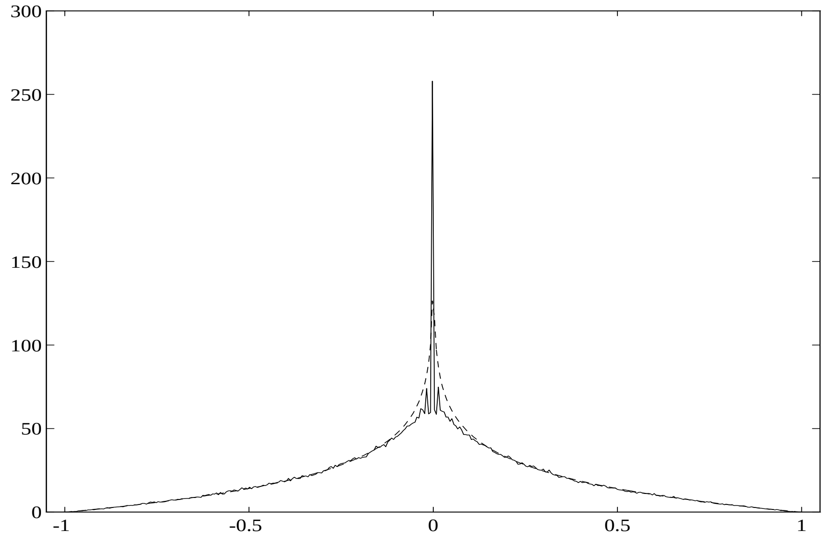

Figure 7.8: A cross section of the function $p(r, \theta)$ along the $\theta = 0$ axis. The solid curve is determined empirically by computing the Cesaro sums. The dashed curve is a graph of the predicted density from our model. The discrepancy near the origin is due to the approximation of the of the complex exponential term in our model with a Gaussian.

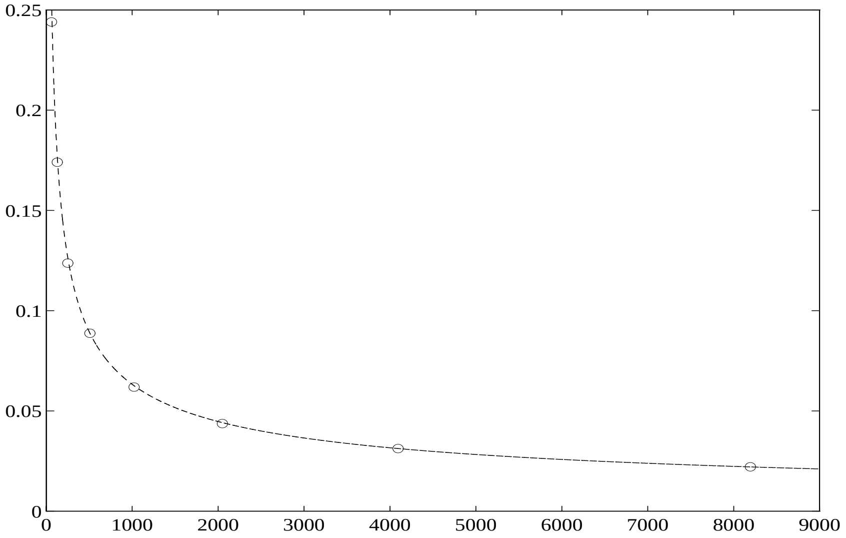

Figure 7.9: Measured versus predicted values of λ_∞ for the Dirac-Fourier dictionary as a function of the dimension N of the space $\mathcal{H}$. The circles correspond to empirically determined values of λ_∞.

The average value of λ_∞ for this equilibrium process is much smaller than the value $\sqrt{\frac{\log N}{N}}$ which would be obtained from a white stationary Gaussian noise. This shows that the realizations of the dictionary noise have energy that is truly well spread over the dictionary elements.

7.5 Asymptotic Evolution of the Residues

Our experiments have shown that residues converge to dictionary noise, and we have characterised this "noise" with a density function derived from a stochastic differential equation model. In this section we modify our stochastic model to obtain information about the time-evolution of the residues close to the attractor. We show that this modified model successfully predicts the macroscopic behavior of the residues for realizations of a Gaussian white noise process.

Our model for the evolution of residues on the attractor in the previous section required two basic assumptions.

1. The correlation ratio on iteration n, λ_n, is approximately equal to a constant.

   We drop this assumption in order to gain information on the convergence to the attractor.

2. The diffusion term A_γ^n in equation (7.23) can be approximated by a random variable with a known magnitude and uniformly distributed random phase which is independent of n.

   This approximation requires first that $\lambda_\infty \ll | < g_\gamma, g_{P^n} > |$ so that we can determine the magnitude of A_γ^n with some degree of accuracy. When λ_∞ is sufficiently small, the phase of A_γ^n is predominantly determined by the phase of $< g_{P^n}, g_\gamma >$. The assumption that this phase is independent of n and uniformly distributed is tantamount to assuming that the energy of the residue is uniformly spread out over the dictionary and that the dictionary elements which are selected do not have any particular correlation that would bias the phase.

We will construct a stochastic differential equation model which predicts the evolution of residues which are close to the attractor in the sense that the energy of the residues is well spread out over the dictionary and the coefficients of the selected dictionary vectors do not have strongly correlated phases. We verify the model numerically for a Gaussian white noise, a process which satisfies the above assumptions.

Let P be a stochastic process such that realizations of P have energy uniformly distributed over the dictionary. We follow the derivation of the model of the previous section with one major change. When we are away from the attractor the correlation ratio λ_n is no longer constant. We replace the constant λ_∞ in our derivation with the variable λ_n. In order to make an approximation of the term A_γ^n of equation (7.23) similar to that of the previous section, we require that $\lambda_n \ll | < g_\gamma, g_{P^n} > |$. Our approximation of the phase of

A^n_γ by a uniformly distributed random phase $e^{i\phi^n_\gamma}$ is justified by our assumption that the selected dictionary elements do not have a correlated structure. We thus obtain

$$< P^{n+1}, g_\gamma >= \begin{cases} \frac{<P^n, g_\gamma>}{\sqrt{1-\lambda_n{}^2}}, & \text{if } < g_\gamma, g_{P^n} >= 0, \\ \sqrt{1-\lambda_n{}^2} < P^n, g_\gamma > + \frac{\lambda_n |<g_\gamma, g_{P^n}>| e^{i\phi^n_\gamma}}{\sqrt{1-\lambda_n{}^2}}, & \text{if } \lambda_n{}^2 \ll |< g_\gamma, g_{P^n} >| \\ 0, & \text{if } g_\gamma = g_{P^n}. \end{cases} \tag{7.40}$$

One example for which all our assumptions hold is the Dirac-Fourier dictionary in an N-dimensional Euclidean space with P equal to a Gaussian white noise process. When $|< g_\gamma, g_{P^n} >|$ is 0 or 1, the update to $< P^n, g_\gamma >$ is exact; otherwise, we know that the energy of the white noise process is uniformly spread over the dictionary vectors, and λ^2_n is $O(\frac{\log N}{N})$ which is much smaller than $|< g_\gamma, g_{P^n} >| = \frac{1}{\sqrt{N}}$, so our approximations are reasonable.

Because we assume that the energy of the residue is uniformly spread over the dictionary we can again make use of the symmetrical positions of the Dirac and Fourier basis vectors. There is an equal probability that the selected vector g_{P^n} belongs to the Dirac or to the Fourier basis. For a fixed g_γ, the first and second updating equations apply with equal frequency, so we average these equations to obtain for $g_\gamma \neq g_{P^n}$ that

$$< P^{n+2}, g_\gamma > - < P^n, g_\gamma >= \frac{\lambda_n e^{i\phi^n_\gamma}}{\sqrt{N}}. \tag{7.41}$$

The energy of the residue is spread over the dictionary elements, so the value of the largest inner product $|< P^n, g_{P^n} >|$ will not be much greater than the value of the next largest inner product. Moreover, λ_n is not large, so the removal of the element g_{P^n} from P^n will not have a large effect on the inner products $|< P^n, g_\gamma >|$. Hence, the value λ_n does not change rapidly. When K is not too large we can approximate the variables $\lambda_n, \lambda_{n+1}, \ldots, \lambda_{n+2K}$ with their mean, $\overline{\lambda}_n$. By the central limit theorem the difference $\frac{1}{2K}(< P^{n+2K}, g_\gamma > - < P^n, g_\gamma >) \approx \frac{\overline{\lambda_n}}{2K\sqrt{N}} \sum_{k=n}^{n+2K} e^{i\phi^k_\gamma}$ tends to a complex Gaussian random variable with variance 1. An important effect of approximating the complex exponential $e^{i\phi^n_\gamma}$ with a Gaussian will be that the model's evolution will be smoother than that of the actual system since we are replacing the exponential term with its running average.

The evolution of the system is thus described by a random walk with a varying step size and a moving absorbing boundary at λ_n. We approximate the difference equation with a continuous time Langevin equation to obtain an explicit solution.

$$\frac{d}{dt} < P^t, g_\gamma >= \frac{\lambda(t)}{2\sqrt{N}} \eta(t), \tag{7.42}$$

where $\eta(t)$ is a complex Weiner process with mean 0 and variance 1. The corresponding Fokker-Planck equation [22] describes the evolution of the probability distribution $p(z, t)$

of $z =< P^t, g_\gamma >$. Since the real and complex parts of $\eta(t)$ have same variance, the solution can be written $p(z,t) = p(r,t)$ where $r = |z|$ and

$$\frac{\partial p(r,t)}{\partial t} = \frac{\lambda(t)^2}{8N} \triangle p(r,t). \tag{7.43}$$

As in the equilibrium case, the boundary condition at 0 will be given by

$$\lim_{r \to 0} \oint_{|z|=r} J \cdot \hat{n} \; d\ell = \frac{1}{2N}, \tag{7.44}$$

where the probability current $J = -\frac{\lambda(t)^2}{8N} \nabla p$. The flow of probability across the boundary at $\lambda(t)$ is driven both by the current and by the motion of the boundary. The flux across $|z| = \lambda(t)$ will be $\frac{1}{2N}$, because we are setting exactly one of the $2N$ inner products to zero each time step. We obtain a relationship between the current at the boundary and the location of the boundary, $\lambda(t)$, using the constraint that the total probability is conserved. We have,

$$\frac{d}{dt} \int_0^{2\pi} \int_0^{\lambda(t)} p(r,t) r dr d\theta = 0, \tag{7.45}$$

from which we obtain,

$$\lambda'(t) \quad = \quad \frac{J(\lambda(t),t) - \frac{1}{4N\pi\lambda(t)}}{p(\lambda(t),t)}. \tag{7.46}$$

We see from 7.43 that the evolution of the probability density function is governed by a heat equation with boundary conditions on moving boundaries. We can think of these equations as governing the depth of a fluid in a cylindrical chamber. On the left end of the chamber there is a small opening and on the right a piston containing an identical opening. The chamber is initially filled with fluid with depth distributed according to $p(r,0)$. The variations in the depth spread out via diffusion. Fluid is pumped into the chamber from the hole on the left at a rate $\frac{1}{4N\pi\lambda(t)}$, and it flows out at an equal rate on the right through the hole in the piston. If the diffusion-induced current at the right edge of the chamber is larger than $\frac{1}{4N\pi\lambda(t)}$, the excess current forces the piston outward. If the induced current is smaller than $\frac{1}{4N\pi\lambda(t)}$, the piston moves inwards to generate additional current. In our numerical experiments the initial wave dies down, and the system evolves to a state of constant flow.

We solve the coupled equations (7.43) and (7.46) using a center-space, forward time finite difference method for $p(r,t)$. The value of $\lambda(t)$ is computed by enforcing the conservation of $\int_0^{\lambda(t)} p(r,t) r dr$. Although the solutions are singular at the origin, our grid values remain bounded because we use a conservation scheme and the integral of this singularity is bounded. We take $\lambda(0)$ to be $E| < P, g_\gamma > |$, which we estimate numerically, and we set $p(r,0)$ equal to a complex Gaussian of mean 0 and variance $\frac{1}{2N}$ truncated at r $=$ $E| < P, g_\gamma > |$ and normalized so that its integral is 1.

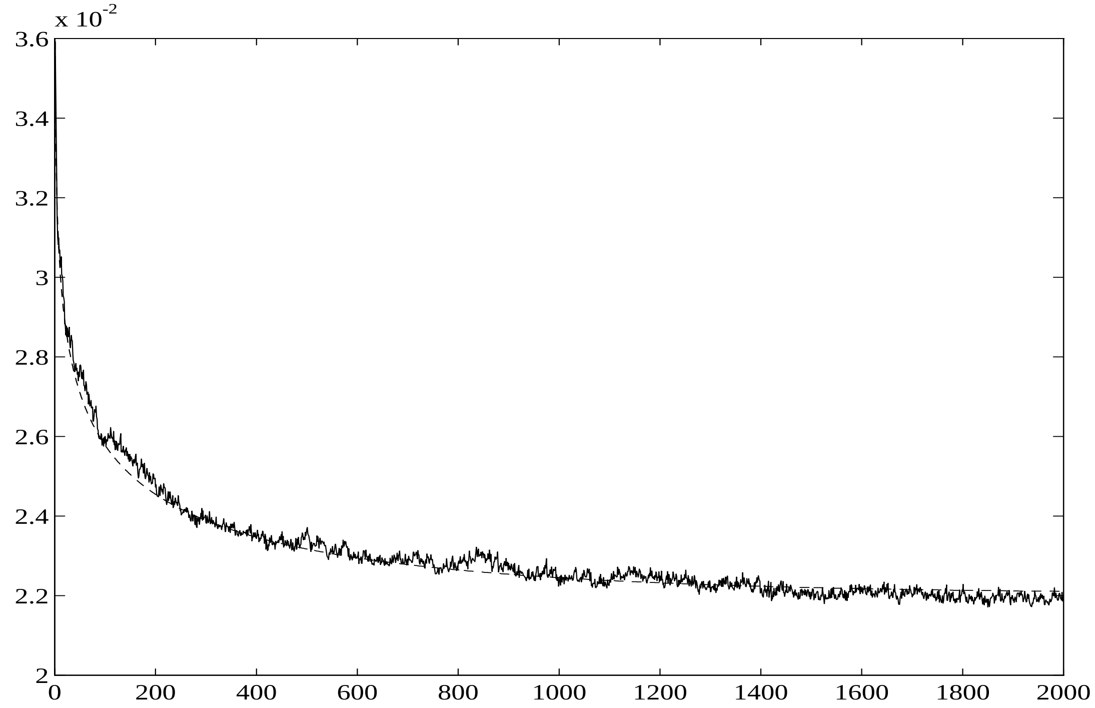

Figure 7.10: The decay of λ as a function of n for a Gaussian white noise with the Dirac-Fourier dictionary. The solid curve is a set of $\lambda(t)$'s obtained from a pursuit, and the dashed curve is a solution $\lambda(t)$ of our model.

Figure 7.10 shows the solution $\lambda(t)$ we obtain for $N = 8192$ together with the values of λ_n obtained from a matching pursuit with a realization of a Gaussian white noise process. Comparing numerical solutions of (7.43) to pursuits with noise, we find that the model provides a good explananation of the macroscopic behavior of λ_n. The measured λ_n has greater variability for several reasons. First, in replacing the $e^{i\phi_\gamma^n}$ term with a Gaussian in our model, we are effectively smooth the evolution. Also, the distribution of the discrete set of inner products $< \tilde{R}^n f, g_\gamma >$ only approximates the smooth density function $p(r,t)$, especially in the tail which determines $\lambda(t)$.

The behavior of the system is consistent with our analysis of section 7.2. The tail of the initial Gaussian represents the part of the signal which is coherent. In the tail of the initial Gaussian both $p(r,t)$ and its slope are close to 0 so $\lambda(t)$ initially decreases very quickly. This corresponds to the small number of random coherent structures present in the white noise being removed. As t increases, the flow from the source starts to pile up at the origin (the number of elements which have small inner product with the residue increases), and the decay of $\lambda(t)$ slows due to the increase in $p(\lambda(t))$ and the decrease in $[-\frac{\lambda^2(t)}{8N}p_r(\lambda(t),t) - \frac{1}{4N\pi\lambda(t)}]$. This corresponds to the energy of the residue spreading out across all of the dictionary elements. The entire system evolves to an asymptotic steady state, which corresponds to the invariant density.

Chapter 8

Learning Optimal Dictionaries

Many types of data, such as recorded speech signals [44], can be modeled as realizations of a random process. We now consider the problem of optimizing a dictionary for decomposing the realizations of a particular random process. Our criterion for optimality is that we minimize the expected error of the M-element expansions,

$$E\|f - \sum_{k=0}^{M-1} \beta_k g_{\gamma_k}\|^2, \tag{8.1}$$

for some fixed M, where f is a realization of our process. We will perform this minimization over a class of dictionaries which is parametrized by a finite parameter set $\mathbf{a} = (a_1, a_2, \ldots, a_K)$. This is a very general class; for instance, any finite dictionary in a finite dimensional space can be represented as such a dictionary. For our numerical experiments, we will optimize a subset of the Gabor dictionary which is parametrized by a single scale.

This type of optimization is particularly important when we work with dictionaries which are characterized by a large number of parameters. For example, [48] [1] expands speech signals into sums of formant-wave functions, a set of waveforms which model the partial response of the vocal tract to a single excitation produced by the vocal cords. The expansion of speech data into sums of these functions allows the extraction of important psychoacoustical parameters such as pitch and the positions of formants. Formant-wave functions are are characterized by translation, modulation, attack rate, and decay rate parameters. If the signal is initially defined over N points, the corresponding dictionary with uniformly sampled parameters contains $O(N^4)$ elements, which is extremely large for speech signals where N is of the order of $10,000$ samples. We need to subsample this dictionary to reduce the computational complexity of the decomposition, but we want to do so in such a way that we have minimal increase in our error. In this section we describe an algorithm for iteratively optimizing an initial set of dictionary parameters which is based on the Lloyd algorithm, an algorithm used for optimizing vector quantizers. For the speech example above, we can parametrize our subsampling and then optimize these subsampling parameters for speech data.

8.1 Vector Quantization and the Lloyd Algorithm

The problem of vector quantization is similar to our problem of compact function representation. Let $\mathbf{X}$ be a vector space and let $\mathbf{B}$ be a set $\{b_i\}$. The set $\mathbf{B}$ is called the *code book*, and its elements are called *code words*. For typical vector quantization applications the cardinality of B is much smaller than that of X. Vector quantization is a two-stage process. Each $x \in X$ is mapped to one of the code words by an *encoder*, a function $C : \mathbf{X} \to \mathbf{B}$. The encoded x is an intermediate representation which can be used for compact storage or transmission. We convert this code word back to a vector in $\mathbf{X}$ with a *decoder*, a function $D : \mathbf{B} \to \mathbf{X}$. We would like to minimize the distortion incurred through this encoding/decoding process for realizations of a random process, i.e. we seek to minimize the expected quantization error

$$E\|D(C(x)) - x\|^2 = E\|\tilde{x} - x\|^2, \tag{8.2}$$

where x is a realization of the process and $\tilde{x}$ is the quantized value of x.

8.1.1 Conditions for Optimality

We assume that our quantizer is memoryless, i.e. that it makes no use of information about what it has encoded in the past. For such a quantizer to be optimal, in the sense that it minimizes (8.2), we have two necessary conditions:

1. *Nearest Neighbor Condition.* Suppose we are given a decoder D such that $D(b_i) = \tilde{x}_i$. Then no encoder can do better than to assign to x the code word b_i which minimizes $\|x - D(b_i)\|$. Equivalently, we quantize to $\tilde{x}_i$ the set of all x for which

$$\|x - \tilde{x}_i\| < \|x - \tilde{x}_j\|, \; i \neq j. \tag{8.3}$$

   The sets $V_i = \{x : \|x - \tilde{x}_i\| < \|x - \tilde{x}_j\|, i \neq j\}$ are called the Voronoi regions, or nearest neighbor cells, of the decoder values $\tilde{x}_i$. The encoder partitions the space $\mathbf{X}$ according to the code words assigned to each $x \in \mathbf{X}$. These code-word partitions are given by the sets $B_i = \{x : C(x) = b_i\}$. The encoder is optimal when the encoder partition B_i for each code word b_i is equal to the Voronoi region of the corresponding decoder value $\tilde{x}_i$.

2. *Generalized Centroid Condition.* Suppose we are given an encoder C. No decoder can do better than to assign to the code word b_i the generalized centroid of the set B_i. The generalized centroid of B is defined to be the $y_i \in \mathbf{X}$ which minimizes $E_{x \in B_i}\|y_i - x\|^2$. The decoder which assigns y_i to b_i, i.e. $D(b_i) = \tilde{x}_i = y_i$, is thus optimal.

A quantizer is uniquely characterized by the partition $\{B_i\}$ of $\mathbf{X}$ and the set of outputs $\{\tilde{x}_i\}$ of the decoder. Given a set $\{\tilde{x}_i\}$, the nearest neighbor condition tells us how to generate an optimal set $\{B_i\}$. Given $\{B_i\}$, the generalized centroid condition tells us how to generate an optimal set $\{\tilde{x}_i\}$.

8.1.2 The Lloyd Algorithm

The Lloyd algorithm proceeds by iterating the following steps on an arbitrary quantizer, $\{\{B_i\}, \{\tilde{x}_i\}\}$.

1. Using the initial decoder, we optimize the encoder according to the nearest neighbor condition. We use the given $\{\tilde{x}_i\}$ to generate an optimal partition $\{B_i'\}$ of $\mathbf{X}$.

2. Using the encoder from the first step, we optimize the decoder according to the generalized centroid condition. We use the optimized partition $\{B_i'\}$ to generate an optimal set of decoder values $\{\tilde{x}_i'\}$.

3. We repeat steps 1 and 2 using the quantizer $\{\{B_i'\}, \{\tilde{x}_i'\}\}$.

Each step either decreases or leaves unchanged the expected error (8.2). The error is nonincreasing and bounded below by 0, so the algorithm produces a sequence of quantizers whose average errors decrease to some lower limit.

Satisfying both the conditions for optimality does not guarantee that a given quantizer is globally optimal, i.e. that there is no other quantizer which gives a lower average error. The average error is a functino on the very high dimensional domain consisting of the vectors associated with the code words, and this error function is in general quite complicated, possessing numerous local extrema. A quantizer with discrete inputs and a finite training set that satisfies the conditions for optimality can be shown to be locally optimal in the sense that small perturbations of the code vectors make the average error worse [27].

8.2 The Lloyd Algorithm for Multistage Quantizers

8.2.1 Optimal Multistage Encoders

We can apply this encoder/decoder framework to dictionary expansions. We take for our dictionary $\mathcal{D} = \{g_\gamma(\mathbf{a})\}_{\gamma\in\Gamma}$, where $\mathbf{a}$ is a set of K parameters. Our goal is to modify these parameters to improve function approximations. Consider the M-element expansion of f into $f = \sum_{k=0}^{M-1} \beta_k g_{\gamma_k}(\mathbf{a}) + R^M f$. This expansion can be viewed as a quantization of f, where a code word consists of the M coefficients β_k and the M indices γ_k.

The function f is encoded to the pair $(\mathcal{G}, \mathcal{B})$, where $\mathcal{B} = (\beta_0, \ldots, \beta_{M-1})$ and $\mathcal{G} = (\gamma_0, \ldots \gamma_{M-1})$. In this chapter we work with encoders for which the coefficients $\mathcal{B}$ can be written explicitly in terms of the function f and the coefficients $\mathcal{G}$. For example, with an orthogonal pursuit, the coefficients $\mathcal{B}$ are given by the orthogonal projection of f onto the subspace spanned by $\{g_\gamma\}_{\gamma\in\mathcal{G}}$. For simplicity, then, we will describe functions as being encoded to the set of indices $\mathcal{G}$.

The decoder assigns to the code word $\mathcal{G}$ the sum

$$D(f,\mathcal{G},\mathbf{a}) = \sum_{\gamma\in\mathcal{G}} \beta_\gamma g_\gamma(\mathbf{a}). \tag{8.4}$$

Unless we specify otherwise, we assume that the values β_γ in (8.4) are the optimal values obtained by projecting f onto the span of $\{g_\gamma\}_{\gamma\in\mathcal{G}}$.

The quantization error is $R^M f$. Our goal is to minimize the expected norm of this quantization error when f is the realization of some random process. The parameter set $\mathbf{a}$ completely specifies the decoder in the above decoding process, so we can use the conditions for optimality defined above to find an optimal encoder corresponding to the decoder with a particular parameter set $\mathbf{a}$. To satisfy condition (8.3) for our expansion, we must encode every f to a set of indices $\mathcal{G}$ which satisfies

$$\|f - D(f,\mathcal{G},\mathbf{a})\| \leq \|f - D(f,\mathcal{G}',\mathbf{a})\|. \tag{8.5}$$

Thus, the notion of an optimal encoder for an multi-stage vector quantization is identical to our M-optimal approximation problem, so obtaining an optimal encoding for a particular parameter set $\mathbf{a}$ requires that we solve the NP-hard optimal approximation problem.

8.2.2 Optimal Multistage Decoders

Suppose that we have an optimal encoder for a given dictionary. We can use this encoder to find an improved decoder using a variation on our above condition for optimality. The optimal encoder partitions the functions in the signal space $\mathcal{H}$ according to the $\mathcal{G}$'s to which they are encoded. Let $B_\mathcal{G}$ be the set of all functions $f \in \mathcal{H}$ which are encoded to the set of indices $\mathcal{G}$. For the decoder to be optimal for this encoding, it must assign to each partition $B_\mathcal{G}$ its generalized centroid, the value $\tilde{f}_\mathcal{G}$ which minimizes $E_{f\in B_\mathcal{G}}\|f - \tilde{f}_\mathcal{G}\|^2$. We cannot in general satisfy this optimality condition. When the dictionary is infinite, we must compute the centroids of an infinite number of partitions. Even for a finite dictionary of size M there are $\begin{pmatrix} M \\ m \end{pmatrix}$ partitions $B_\mathcal{G}$, so if M is large we must compute an inordinately large number of generalized centroids. We would like to maintain the structure of the decoder in (8.4), and restrict our dictionary changes to modifications of the parameter set $\mathbf{a}$. Even if we are able to find the generalized centroids of each region $B_\mathcal{G}$, the K parameters must satisfy $\begin{pmatrix} M \\ m \end{pmatrix}$ constraints.

We can use a nonlinear minimization technique to modify the parameters $\mathbf{a}$ in order to reduce the expected global quantization error $E\|f - D(f,\mathcal{G},\mathbf{a})\|^2$ rather than trying to reduce the error from each partition. If we are then able to find an optimal encoder for any given decoder, we can perform an analog of the Lloyd algorithm.

We set the parameters $\mathbf{a}$ to some initial value and then optimize $\mathbf{a}$ for a training set $\{f^{(k)}\}_{1\leq k\leq n}$ by the following procedure.

1. For each $f^{(k)}$ in the training set, we find an optimal encoding $\mathcal{G}^{(k)}(\mathbf{a})$ so that
$$\|f^{(k)} - D(f^{(k)}, \mathcal{G}^{(k)}(\mathbf{a}), \mathbf{a})\|^2 \leq \|f^{(k)} - D(f^{(k)}, \mathcal{G}', \mathbf{a})\|^2 \tag{8.6}$$
for all other encodings $\mathcal{G}'$.

2. Given this encoding, we use a minimization step to improve the dictionary. We find an $\mathbf{a}'$ so that
$$\frac{1}{n}\sum_k \|f^{(k)} - D(f^{(k)}, \mathcal{G}^{(k)}(\mathbf{a}), \mathbf{a}')\|^2 < \frac{1}{n}\sum_k \|f^{(k)} - D(f^{(k)}, \mathcal{G}^{(k)}(\mathbf{a}), \mathbf{a})\|^2. \tag{8.7}$$
If no such $\mathbf{a}'$ can be found we terminate the algorithm. Otherwise, we set $\mathbf{a} = \mathbf{a}'$ and repeat.

After the algorithm finishes, we perform one final step of the encoding so that the encoder and the decoder use the same value of $\mathbf{a}$. Each time step 2 is followed by step 1, we reduces or leave unchanged the global error $\frac{1}{n}\sum_k \|f^{(k)} - D(f^{(k)}, \mathcal{G}, \mathbf{a})\|^2$, so the algorithm generates a sequence of quantizers whose average error at step m, $e(\mathbf{a}_m)$ decreases to some lower limit $e(a_\infty)$. By iterating the algorithm a sufficient number of times, we can generate a parameter set $\mathbf{a}$ which gives an error which is arbitrarily close to $e(\mathbf{a}_\infty)$. In our numerical implementation, we stop the algorithm when the fractional decrease in the error, $\frac{(e(\mathbf{a}_m)-e(\mathbf{a}_{m+1}))}{e(\mathbf{a}_m)}$ is less than some threshold.

8.3 Optimizing Dictionaries with the Levenberg-Marquardt Algorithm

We now describe in more detail the process of optimizing the parameter values of the dictionary elements. We suppose for the moment that we have an optimal encoding available. We denote by $S_\mathcal{G}(\mathbf{a})$ the operator which maps an M-vector of coefficients $(\beta_0, \ldots, \beta_{M-1})$ to the sum $\sum_{k=0}^{M-1} \beta_k g_{\gamma_k}(\mathbf{a})$. If we express the elements g_{γ_k} as vectors, $S_\mathcal{G}(\mathbf{a})$ will be the matrix whose columns are the dictionary elements with indices in $\mathcal{G}$. We assume each g_γ is a C^2 function of the parameters $\mathbf{a}$, and that $\|g_\gamma(\mathbf{a})\| = 1$ for all values of γ and $\mathbf{a}$.

The optimal set of coefficients $\mathcal{B}$ for an expansion of f over $\{g_\gamma\}_{\gamma\in\mathcal{G}}$ is given by the coefficients of the projection of f onto the span of the $\{g_\gamma\}_{\gamma\in\mathcal{G}}$,
$$\mathcal{B} = (S^*_\mathcal{G}(\mathbf{a})S_\mathcal{G}(\mathbf{a}))^{-1}S^*_\mathcal{G}(\mathbf{a})f. \tag{8.8}$$
The quantization error will be
$$\begin{aligned} \|R^M f\| &= \|f - S_\mathcal{G}(\mathbf{a})\mathcal{B}\| & (8.9)\\ &= \|f - S_\mathcal{G}(\mathbf{a})(S^*_\mathcal{G}(\mathbf{a})S_\mathcal{G}(\mathbf{a}))^{-1}S^*_\mathcal{G}(\mathbf{a})f\|. & (8.10)\end{aligned}$$

We estimate the expected error $E\|f - D(f, \mathcal{G}, \mathbf{a})\|^2 = E\|R^M f\|^2$ by taking a set of n realizations $f^{(k)}$ of the process and computing the sum

$$E\|f - D(f, \mathcal{G}, \mathbf{a})\| \approx e(\mathbf{a}) = \frac{1}{n}\sum_{k=1}^{n} \|R^M f^{(k)}\|^2 \tag{8.11}$$

Using (8.9) we can express $e(\mathbf{a})$ as a differentiable function of $\mathbf{a}$ and use the Levenberg-Marquardt algorithm to minimize this error by iteratively modifying the parameters $\mathbf{a}$. If we are far from a minimum, we use a steepest descent method to minimize (8.11). At each iteration we compute an updated set of parameters $\mathbf{a}'$,

$$\mathbf{a}' = \mathbf{a} - C\nabla e(\mathbf{a}), \tag{8.12}$$

for some constant C. If we are sufficiently close to a minimum and if we can compute the Hessian accurately, we can reduce (8.11) using Newton's method. We take

$$\mathbf{a}' = \mathbf{a} - H^{-1}\nabla e(\mathbf{a}). \tag{8.13}$$

The Levenberg-Marquardt method adaptively interpolates between these two methods by making use of the following observations [45][40]. Let a_i and a_j be any two single parameters of $\mathbf{a}$. We have

$$\frac{\partial e(\mathbf{a})}{\partial a_i} = \frac{1}{n}\sum_k 2Re < \frac{\partial R^m f^{(k)}}{\partial a_i}, R^m f^{(k)} > \tag{8.14}$$

and

$$\frac{\partial^2 e(\mathbf{a})}{\partial a_i \partial a_j} = \frac{1}{n}\sum_k 2Re < \frac{\partial R^m f^{(k)}}{\partial a_i}, \frac{\partial R^m f^{(k)}}{\partial a_j} > + \frac{1}{n}\sum_k 2Re < R^m f^{(k)}, \frac{\partial^2 R^m f^{(k)}}{\partial a_i \partial a_j} > . \tag{8.15}$$

We assume that the residues $R^m f^{(k)}$ are small, so we can approximate the second derivative (8.15) with just the first sum of first derivative terms. We use this approximation for the second derivative to form an approximation for the Hessian in (8.13).

Dimensional analysis of the steepest descent algorithm indicates that we can estimate the order of magnitude of the requisite step from the reciprocals of the diagonal elements of the Hessian. In our above approximation for the Hessian, these diagonal elements are given by the positive quantities $\frac{2}{n}\sum_k \|\frac{\partial R^m f^{(k)}}{\partial a_i}\|^2$, so in using these to determine the step size, we do not step against the gradient.

For Newton's method, we have

$$H\delta\mathbf{a} = -\nabla e(\mathbf{a}), \tag{8.16}$$

where H is our approximation to the Hessian, and for the steepest descent, we now have

$$\lambda D\delta\mathbf{a} = -\nabla e(\mathbf{a}), \tag{8.17}$$

where λ is a scaling constant and D is the diagonal matrix with the diagonal equal to that of the approximated Hessian. The Levenberg-Marquardt algorithm combines these two equations, and obtains successive values of the $\mathbf{a}$ by solving,

$$(H + \lambda D)\delta\mathbf{a} = -\nabla e(\mathbf{a}). \tag{8.18}$$

When λ is large, we are effectively performing a small step of steepest descent. When λ is small, we are performing a step of Newton's method. The value of λ is adjusted each iteration. If a step causes the error $e(\mathbf{a})$ to increase, the step is discarded and λ is increased so that the next iteration will be more like a steepest descent step with a small stepsize. If a step causes the error to decrease, λ is decreased so that the next iteration will be more like a step of Newton's method.

8.4 Implementation

Computing the optimal encoding is too slow in practice because the problem is NP-complete. We instead approximate the optimal encoding using a pursuit algorithm. When we use a non-optimal encoder, our algorithm does not necessarily yield decreasing values of $e(\mathbf{a}_m)$ as m increases. The reason is that after we have changed the parameter $\mathbf{a}$ to $\mathbf{a}'$ to decrease (8.7), the new average error obtained from the modified decoder acting on the modified encoding,

$$\frac{1}{n}\sum_k ||f^{(k)} - D(f^{(k)}, \mathcal{G}^{(k)}(\mathbf{a}'), \mathbf{a}')||^2,$$

is not necessarily less than the average error obtained from the modified decoder operating on the old encoding

$$\frac{1}{n}\sum_k ||f^{(k)} - D(f^{(k)}, \mathcal{G}^{(k)}(\mathbf{a}), \mathbf{a}')||^2.$$

When the change in the parameter value $\delta\mathbf{a}$ is sufficiently small, the change of $\mathbf{a}$ does not affect the encodings $\mathcal{G}^{(k)}(\mathbf{a})$ (unless one of the functions $f^{(k)}$ lies on a boundary of one of the nearest neighbor cells $B_{\mathcal{G}}$, a set of measure 0), so the algorithm still works. When $\delta\mathbf{a}$ is large enough that a few of the $f^{(k)}$'s have different encodings after the modification of $\mathbf{a}$, the algorithm will still cause $e(\mathbf{a})$ to decrease as long as any increase in the error due to the non-optimal repartitioning is offset by the decrease in the error from adjusting $\mathbf{a}$.

Thus, when the change in $\mathbf{a}$ is sufficiently small, our algorithm will work with a non-optimal encoder. The Levenberg-Marquardt method discards steps for which $e(\mathbf{a}')$ increases, so by using such an adaptive minimization scheme we are assured that the step size $\delta\mathbf{a}$ will be small enough that we obtain a sequence of $\mathbf{a}_m$ for which the average error decreases.

When we replace the encoder with a matching pursuit in our algorithm, the coefficients $\mathcal{B}$ are no longer obtained by projecting f onto the span of $\{g_\gamma\}_{\gamma\in\mathcal{G}}$, so we will have to modify our expression (8.8) for obtaining $\mathcal{B}$ from f and $\mathcal{G}$. We first obtain an explicit expression for the quantization error $R^M f$ as a product of the projection operators,

$$R^M f = \left[\prod_{\gamma \in \mathcal{G}} (I - g_\gamma g_\gamma^*) \right] f. \tag{8.19}$$

We can simplify the form of the error by making the assumption that the first M elements of the optimal expansion form a nearly orthogonal set, when M is not too large. This assumption clearly holds for $M = 1$. When the energy of the functions being decomposed is spread uniformly over several dictionary elements, the pursuit will select first the elements whose inner product with elements previously selected is small. There is a bias against selecting an element g_γ which correlates with previously selected ones because in removing the element g_{γ_k} which correlates with g_γ, the pursuit decreases the inner product $| < R^k f, g_\gamma > |$. We thus assume that the inner products $< g_{\gamma_i}, g_{\gamma_j} >= O(\epsilon)$ when $i \neq j$. Expanding the series (8.19) and retaining first order terms in ϵ gives

$$\|R^M f\|^2 = \|f\|^2 - \sum_{0 \leq j < M} |< f, g_{\gamma_j} >|^2 + \sum_{0 \leq j < M} \sum_{k \neq j} < f, g_{\gamma_j} >< g_{\gamma_j}, g_{\gamma_k} >< g_{\gamma_k}, f > + O(\epsilon^2). \tag{8.20}$$

This representation allows us to compute quickly the derivatives with respect to the parameters $\mathbf{a}$ that we need for the Levenberg-Marquardt minimization procedure.

We can make a similar simplification when approximating the optimal encoder with an orthogonal pursuit. We again assume that $< g_{\gamma_i}, g_{\gamma_j} >= O(\epsilon)$ when $i \neq j$. The matrix $S_{\mathcal{G}}^* S_{\mathcal{G}}$ is very close to the M by M identity matrix. We can write $S_{\mathcal{G}}^* S_{\mathcal{G}}$ as $I + A$ where A has small norm. We can approximate

$$\begin{aligned} (S_{\mathcal{G}}^* S_{\mathcal{G}})^{-1} &= (I + A)^{-1} \\ &= I - A + O(A^2) \\ &\approx 2I - S_{\mathcal{G}}^* S_{\mathcal{G}}. \end{aligned} \tag{8.21}$$

8.5 Results

We test our algorithm by adapting a subset of the Gabor dictionary to a process which generates random chirps. Our dictionary consists of the Gabor functions $\frac{1}{\sqrt{s_0}} g(\frac{t-u}{s_0}) e^{i\xi t}$ periodized on the domain $t \in [0, 2\pi)$. Here $g(t)$ is the Gaussian $2^{1/4} e^{-\pi t^2}$. The dictionary contains elements for all values u and ξ in $[0, 2\pi)$. The scale s_0 is fixed. Our source generates chirps of the form $Ce^{ia(t-b)^2}$, periodized on the domain, where the value of b is uniformly randomly distributed in $[0, 2\pi)$.

The chirps generated by this source have support in the time-frequency plane on lines with slope $2a$ and time-axis intercept $2ab$. We can envision the expansions of these chirps over the fixed-scale Gabor elements as a covering of this slanted line with small rectangles.

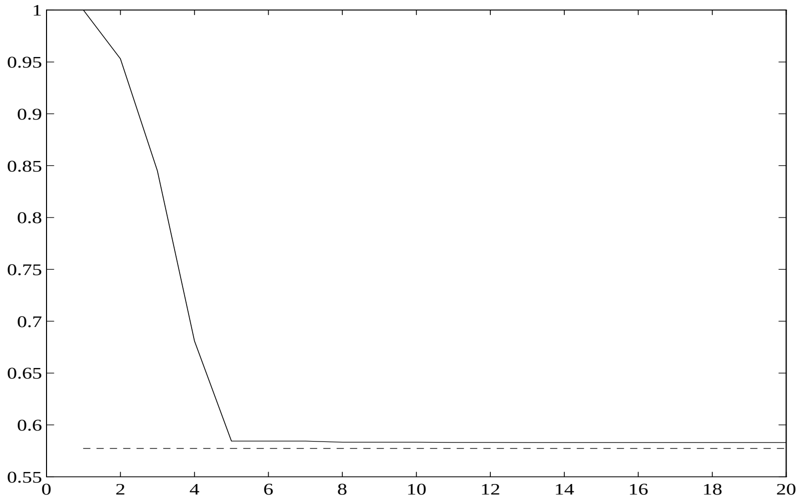

Figure 8.1: Evolution of the scale parameter s_0 under iterations of the adaptation algorithm. The vertical axis is the scale and the horizontal the iteration of the algorithm. The dotted line corresponds to $s_0 = s_{max}$.

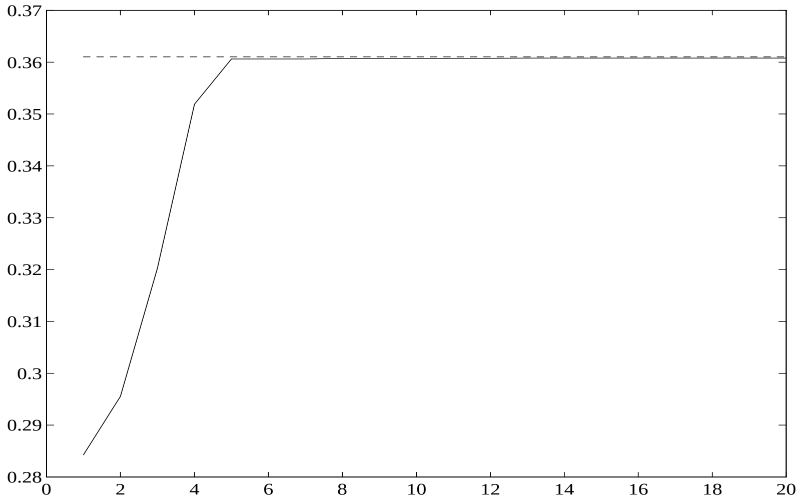

Figure 8.2: Evolution of the average residual norm $||R^M f||$ for the training set under iterations of the adaptation algorithm. The vertical axis is the error and the horizontal the iteration of the algorithm. The dotted line is the average residual norm when $s_0 = s_{max}$.

The eccentricity of the rectangles is controlled by the scale parameter. The inner product of one of the dictionary elements and one of these chirps is maximal when $\xi = 2a(u-b)$ and $s_0 = \sqrt{\frac{\pi}{a}}$, i.e. when the diagonal of the rectangles in the time-frequency is aligned with the chirp. If we look at the time-frequency energy representations of the expansions of one of these chirps, we see that in the time-frequency plane the energy of the selected atoms line up along the $\xi = 2a(u-b)$, and that there is little energy overlap among the largest selected elements. We expect, then, that we will obtain an optimal decomposition when $s_0 \approx s_{\mathrm{max}} = \sqrt{\frac{\pi}{a}}$.

For our numerical experiments we take $a = 9$. We find that the parameter s_0 quickly converges to a value close to $\sqrt{\frac{\pi}{a}}$. For our experiments we used a training set of 20 test vectors in a space of dimension 64, and we took $M = 10$. Figures 8.1 and 8.2 show the evolution of the parameter s_0 and the average error $e(s_0)$ over a number of iterations of the optimization algorithm. We see, then, that the use of the non-optimal pursuit does not cause problems the convergence of the algorithm.

We note that the algorithm converges to a value which is slightly different from the value $\frac{\pi}{a}$ which we expect to be optimal, and when we compute the average error from the training set with the value $s_0 = \frac{\pi}{a}$, we find that it is less than the error obtained from the value to which the algorithm has converged. In implementing our algorithm with more complicated parameter sets, we will need to incorporate some form of stochastic relaxation into our minimization step, such as simulated annealing [37] [41] to avoid becoming trapped in the numerous local minima of the error surface.

Chapter 9

Conclusion

The problem of optimally approximating a function with a linear expansion over a redundant set is a computationally intractable one. The greedy matching pursuit algorithms provide a means of quickly computing compact approximations. The orthogonalized matching pursuit algorithm converges in a finite number of steps in finite dimensional spaces. The much faster non-orthogonal matching pursuits yield comparable expansions for the coherent portion of the signal.

Renormalized matching pursuits possess local topological properties like those of chaotic maps, including local separation of points, and local mixing of the domain. For a particular dictionary, the renormalized pursuit is in fact chaotic and ergodic. Ergodic pursuits possess invariant measures from which we obtain a statistical description of the residues.

For dictionaries which are invariant under the action of a group operator, we can construct a choice function which preserves this invariance. We can deduce properties of the invariant measure of a pursuit with such a dictionary; in particular, the invariant density function of a translation and modulation invariant pursuit will be stationary and white.

Numerical experiments with the Dirac-Fourier dictionary show that the asymptotic residues of the pursuit converge to dictionary noise, the realizations of a white, stationary process. Our stochastic differential equation model shows that the coherent structures, the elements with large inner products $|<R^n f, g_\gamma>|$, are quickly removed by the pursuit. The asymptotic convergence rate is slow, and the asymptotic inner products $<R^n f, g_\gamma>$ essentially perform a random walk until they reach a constant λ_∞ and are selected.

With an appropriate dictionary, the expansion of a signal into its coherent structures provides a close approximation with a small number of terms. We can adapt a dictionary for decomposing a given class of signals using a variant of the generalized Lloyd algorithm.

An important area for further research is the problem of extracting higher level features from signal decompositions. Given an expansion of a speech signal into Gabor functions, for example, can we more efficiently extract phonemes? By adopting a hierarchical structure, like that of the cortex, it may be possible to extend the pursuit algorithm efficiently to the extraction of higher level features.

Another area for further refinement is the optimality criterion used for the approximations themselves. As we have shown, the current minimal approximation error criterion leads to instabilities in the expansions and is partially responsible for the intractability of the optimal approximation problem. A modification of the optimality criterion could lead to more stable expansions and more efficient algorithms.

Bibliography

[1] C. d'Alessandro, "Time-frequency modifications using an elementary waveform speech model," Proceedings of ESCA-Eurospeech 89, Paris, October 1989.

[2] J. Banks, J. Brooks, G. Cairns, G. Davis, and P. Stacey, "On Devaney's Definition of Chaos," *American Mathematical Monthly*, Vol. 99, No. 4, 332-334. April 1992.

[3] H. B. Barlow, "Single units and sensation: A neuron doctrine for perceptual psychology" *Perception*, vol. 1, 371-394, 1972.

[4] H. B. Barlow and Peter Földiák, "Adaptation and Decorrelation in the Cortex." *The Computing Neuron*, Durbin, Miall, and Mitchison, eds.

[5] S. A. Billings, M. J. Korenberg, and S. Chen, "Identification of non-linear optput-affine systems using an orthogonal least-squares algorithm," *International Journal of Systems Science*, Vol. 19, 1559-1568.

[6] A. Buzo, A. H. Gray, R. M. Gray, and J. D. Markel. Speech coding based on vector quantization. *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-28:562-574, October 1980.

[7] S. Chen, S. A. Billings, and W. Luo, "Orthogonal least squares methods and their application to non-linear system identification," *International Journal of Control*, vol. 50, No. 5, 1873-1896, 1989.

[8] L. Cohen, "Time-frequency distributions: a review" Proceedings of the IEEE, Vol. 77, No. 7, 941-979, July 1989.

[9] R. Coifman and V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 36:961-1005, September 1990.

[10] P. Collet and J.P. Eckmann. *Iterated Maps on the Interval as Dynamical Systems*, Birkhauser, Boston, 1980.

[11] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*, McGraw-Hill, New York, 1991.

[12] V. Chvatal, "A greedy heuristic for the set-covering problem," *Mathematics of Operations Research*, Vol. 4, No. 3, 233-235.

[13] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Series in Appl. Math., SIAM, 1991.

[14] R. L. Devaney, *An Introduction to Chaotic Dynamical Systems*, Addison-Wesley Publishing Company, Inc., New York, 1989.

[15] R. A. DeVore, B. Jawerth, V. Popov, "Compression of wavelet decompositions," *American Journal of Mathematics*, 114 (1992): 737-785.

[16] R. A. DeVore, B. Jawerth, B. J. Lucier, "Image compression through wavelet transform coding," *IEEE Trans. Info. Theory*, Vol. 38, No. 2, 719-746, March 1992.

[17] R. A. DeVore, B. Jawerth, B. J. Lucier, "Surface compression," *Computer Aided Geometric Design*, Vol. 9, 1992, 219-239.

[18] D. L. Donoho, "Wavelet shrinkage and W.V.D.: a 10-minute tour," *Progress in Wavelet Analysis and Applications*, Y. Meyer and S. Roques (eds.), Editions Frontieres, Gif-sur-Yvette, France, 1993, 109-128.

[19] R. J. Duffin and A. C. Schaeffer, "A class of nonharmonic Fourier series," *Trans. Amer. Math. Soc.*, 72, pp. 341-366.

[20] J. H. Friedman and W. Stuetzle, "Projection pursuit regression," Journal of the American Statistical Association, Vol. 76, pp. 817-823, 1981.

[21] J. H. Friedman and J. W. Tukey, "A projection pursuit algorithm for exploratory data analysis," *IEEE Trans. Comput.,* C-23:881-889, 1970.

[22] C. W. Gardiner, *Handbook of Stochastic Methods,* Springer-Verlag, New York, 1985.

[23] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman and Co., New York, 1979.

[24] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Boston, 1992.

[25] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1989.

[26] R. Gray, "Vector quantization", *IEEE Acoustic Speech and Signal Processing Magazine*, April 1984.

[27] , "Locally optimal block quantizer design," *Inform. and Control*, 45:178-198, May 1980.

[28] P. R. Halmos, *Lectures on Ergodic Theory*, The Mathematical Society of Japan, Tokyo, 1956.

[29] F. C. Hoppensteadt, *Analysis and Simulation of Chaotic Systems*, Springer-Verlag, New York, 1993.

[30] P. J. Huber, "Projection Pursuit," *The Annals of Statistics*, vol. 13, No. 2, p. 435-475, 1985.

[31] D. Johnson, "Approximation algorithms for combinatorial problems," *J. Comput. System Sci.*, 9:256-278.

[32] J. B. Kruskal, "Toward a practical method to help uncover the structure of a set of multivariate observations by finding the linear transformation which optimizes a new 'index of condensation."' in R. C. Milton and J. A. Nelder, eds., *Statistical Computation.* Academic Press, New York, 1969.

[33] A. Lasota and J. A. Yorke, "On the existence of invariant measures for piecewise monotonic transformations," *Transactions of the American Mathematical Society*, 186 (1983), 481-488.

[34] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inform. Theory*, IT-28:127-135, March 1982.

[35] L. Lovász, "On the ratio of optimal integral and fractional covers," *Discrete Math.* 13:383-390.

[36] L. K. Jones, "On a conjecture of Huber concerning the convergence of projection pursuit regression", *The Annals of Statistics*, vol. 15, No. 2, p. 880-882, 1987.

[37] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, 220:219-227, 1983.

[38] A. Lasota and M. Mackey, *Probabilistic Properties of Deterministic Systems,* Cambridge University Press, New York, 1985.

[39] S. Mallat and Z. Zhang "Matching Pursuit with Time-Frequency Dictionaries", *IEEE Trans. on Signal Processing*, Dec. 1993.

[40] D. W. Marquardt, *SIAM Journal*, Vol. 11, pp. 431-441.

[41] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *J. Phys. Chem,* 21:1087, 1953.

[42] Y. Meyer (translated by R. Ryan), *Wavelets: Algorithms and Applications*, SIAM, Philadelphia, 1993.

[43] Y. C. Pati R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decomposition," *Proceedings of the* 27^{th} *Annual Asilomar Conference on Signals, Systems, and Computers*, Nov. 1993.

[44] T. Parsons, *Voice and Speech Processing*, McGraw Hill, New York.

[45] W. H. Press, S. A. Teukolsky, W. T. Vettering, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, New York, 1992.

[46] S. Qian and D. Chen, "Signal Representation via Adaptive Normalized Gaussian Functions," *IEEE Trans. on Signal Processing*, vol. 36, no. 1, Jan. 1994.

[47] M. Reed and B. Simon, *Methods of Modern Mathematical Statistics, Vol. 1*, Academic Press, New York, 1972.

[48] X. Rodet, "Time-Domain Formant-Wave-Function Synthesis," *Computer Music Journal*, vol. 8, no. 3, 1985.

[49] H. L. Royden, *Real Analysis*, Macmillan Publishing Company, New York, 1988.

[50] M. J. Sabin and R. M. Gray. Global convergence and empirical consistency of the generalized Lloyd algorithm. *IEEE Trans. Inform. Theory,* IT-32:148-155, March 1986.

[51] N. Saito, "Simultaneous Noise Suppression and Signal Compression using a Library of Orthonormal Bases and the Minimum Description Length Criterion," *Wavelets in Geophysics,* to appear.

[52] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, 27:379-423, 623-656, 1948.

[53] L. Blum, M. Shub, and S. Smale, "On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions, and universal machines," *Bulletin of the American Mathematical Society*, Vol. 21, No. 1, 1-46, July 1989.

[54] P. Switzer, "Numerical classification," in *Geostatistics*, Plenum, New York, 1970.

[55] , "Numerical classification applied to certain Jamaican Eocene nummulitids," *Math. Geol.* 3:297-311.

[56] P. Walters, *Ergodic theory–Introductory Lectures*, Springer-Verlag, New York, 1975.

[57] R. M. Young, *An Introduction to Nonharmonic Fourier Series,* Academic Press, New York, 1980.

[58] Z. Zhang, "Matching Pursuit," Ph.D. dissertation, New York University, 1993.