

# C20220 Introduction

Daniel Richardson, Norwood House 6.29, email [masdr@bath.ac.uk](mailto:masdr@bath.ac.uk)

April 23, 2009

Welcome to Fundamentals of Pattern Analysis. I hope that you enjoy the course.

Office hours, to be arranged, will be on Tuesday and Thursday. There will usually be two lectures per week, and lab sessions have also been booked. There will be tutors in the labs to help you get started, and to answer your questions, and to expand on some of the topics in the course. We will be using Maple and Matlab. Information about the course, notes, and links can be found at my website

`www.bath.ac.uk/~masdr`

The course is in four sections.

1. Fourier transform, and approximation. Applications in music and digital signal processing and image analysis (for example face recognition, medical diagnosis).
2. Probability and Statistics. Bayesian inference. Reasoning with uncertainty.
3. Neural nets and Genetic algorithm, and related subjects. The idea here is to attempt to imitate some of the problem solving techniques used by living things.
4. Principal component analysis, and related subjects.

The overall intention of the course is to understand that part of contemporary mathematics which is most useful for computer science. To understand this course, you will need to have or get hold of some understanding of probability and statistics, calculus, complex numbers, and linear algebra. I recommend John Power's set of notes for CM10197. If you do not have these notes, you can find a link to the Nov 30 2008 version from the cm20220 section of my website. You are strongly advised to review these notes. If you find something in these notes that you do not understand very well, please let me know.

# 1 Approximation and Fourier transforms

We will begin with two problems.

1. The vector approximation problem in  $\mathbf{R}^N$ . Given vectors  $X_1, \dots, X_k$  and  $Y$  in  $\mathbf{R}^N$ , find real coefficients  $c_1, \dots, c_k$  so that  $c_1X_1 + \dots + c_kX_k$  is as close as possible to  $Y$ .
2. The audio approximation problem. Given a continuous function  $f : [-\pi, \pi] \rightarrow \mathbf{R}$ , find real coefficients  $a_0, a_1, b_1, a_2, b_2, \dots, a_k, b_k$  so that the function
$$F_k(t) = a_0/2 + a_1\cos(t) + b_1\sin(t) + \dots + a_k\cos(kt) + b_k\sin(kt)$$
is as close as possible to  $f(t)$  on the interval  $[-\pi, \pi]$ .

Notation. If  $a$  and  $b$  are real numbers,  $[a, b]$  is the interval  $\{x : a \leq x \leq b\}$ .  $F_k(t)$  is called the order  $k$  Fourier approximation to the function  $f(t)$  on the interval  $[-\pi, \pi]$ . As usual, we may write  $F_k(t)$  as  $a_0/2 + \sum_{n=1}^k a_n\cos(nt) + b_n\sin(nt)$ .

For the second problem, we are supposed to think of the function  $f$  as representing a piece of music (or a sound, such as for example a dog howling, or a recording of a voice) over a time interval which has been transformed to  $[-\pi, \pi]$ .

The following important definition connects the two problems.

**Definition 1** A vector space over  $\mathbf{R}$  consists of a set  $V$  together with two operations

addition  $+$  :  $V \times V \rightarrow V$

scalar multiplication :  $\mathbf{R} \times V \rightarrow V$

so that the following axioms are satisfied.

1.  $(\forall X, Y \in V) X + Y \in V$   
 $(\forall r \in \mathbf{R}, X \in V) rX \in V$
2. There is a zero vector in  $V$ , denoted  $0$ , so that  $(V, +, 0)$  is a commutative group.
3.  $c(X + Y) = cX + cY$   
 $c(dX) = (cd)X$   
 $(c + d)X = cX + dX$   
 $1X = X$   
for all  $X, Y \in V$  and for all  $c, d \in \mathbf{R}$ .

**Definition 2** Suppose  $a, b$  are real numbers.

Let  $C[a, b] = \{\text{continuous } f : [a, b] \rightarrow \mathbf{R}\}$

$C[a, b]$  is a very large set, including as it does all possible sounds over time interval  $[a, b]$ . To make it into a vector space, we need to define addition and scalar multiplication.

**Definition 3** Suppose  $f$  and  $g$  are in  $C[a, b]$ . Then  $h = f + g$  is defined as follows.

For all  $t \in [a, b]$ , let  $h(t) = f(t) + g(t)$

Suppose  $r \in \mathbf{R}$ . Then  $h = rf$  is defined as follows.

For all  $t \in [a, b]$ , let  $h(t) = rf(t)$

**Exercise 1** Check that with these definitions in place,  $C[a, b]$  is a vector space over  $\mathbf{R}$ .

In any vector space, we have a collection of common ideas: span, linear combination, linear independence, subspace, basis. See J. Power's notes for definitions in  $\mathbf{R}^N$ , and generalise.

I don't find the definition of subspace in JP's notes, so here it is.

**Definition 4** Let  $V$  be a vector space over the reals, and suppose  $S \subseteq V$ . Then  $S$  is a subspace of  $V$  if

1. S1. Whenever  $X$  and  $Y$  are in  $S$ , so is  $X + Y$
2. S2. For any real number  $r$  and any  $X \in S$ , the scalar product  $rX$  is also in  $S$ .

## 1.1 Dot product in $\mathbf{R}^N$

This should be revision. Suppose  $X = (x_1, \dots, x_N)$  and  $Y = (y_1, \dots, y_N)$  are in  $\mathbf{R}^N$ . Then the dot product (or scalar product) of  $X$  and  $Y$  is defined as follows.

$$X \circ Y = x_1 y_1 + \dots + x_N y_N$$

From this, we get some useful definitions.

For vectors  $X, Y$  in  $\mathbf{R}^N$ ,

1.  $|X| = (X \circ X)^{1/2}$ .  $|X|$  is called the norm (or the size or the length or the magnitude) of  $X$ .
2. The distance between  $X$  and  $Y$  is defined as follows.  $d(X, Y) = |X - Y|$ .
3.  $X$  and  $Y$  are orthogonal, written  $X \perp Y$ , if  $X \circ Y = 0$
4. The angle  $\theta$  between two vectors  $X$  and  $Y$  is defined by  $X \circ Y = |X| |Y| \cos(\theta)$
5.  $X$  is a unit vector if  $|X| = 1$ .

We intend to extend all these ideas to the functions in  $C[a, b]$ . (!) It will turn out, for example that  $\sin(t)$  and  $\cos(t)$  are, in a sense, at right angles to each other over  $[-\pi, \pi]$ .

## 1.2 Inner product for a vector space

We need something which behaves like a dot product for  $C[a, b]$ . This will be called an inner product. The inner product of two functions  $f$  and  $g$  will be written  $\langle f, g \rangle$ . The axioms for an inner product in a vector space  $V$  are the following.

1. ip1) For any  $f, g \in V$ , the inner product  $\langle f, g \rangle$  is a real number.
2. ip2)  $\langle f, g \rangle = \langle g, f \rangle$ , for any  $f, g \in V$
3. ip3)  $\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle$ , for any  $f, g, h \in V$
4. ip4)  $\langle rf, g \rangle = r \langle f, g \rangle$ , for any  $f, g \in V$  and any  $r \in \mathbf{R}$ .
5. ip5)  $\langle f, f \rangle \geq 0$  and  $\langle f, f \rangle = 0$  iff  $f = 0$ , for any  $f \in V$ .

We are supposed to be thinking that  $V$  could be  $C[a, b]$  and  $f, g, h$  could be functions. However,

**Exercise 2** Check that these axioms are satisfied for  $V = \mathbf{R}^N$ , and  $\langle X, Y \rangle = X \circ Y$ .

**Definition 5** An inner product space is a vector space  $V$  over  $\mathbf{R}$  together with an inner product  $\langle, \rangle$  satisfying the above axioms.