MA 751
Part 3

Infinite Dimensional Vector Spaces

1. Motivation: Statistical machine learning and reproducing kernel Hilbert Spaces

Gene expression experiments

Question: Gene expression - when is the DNA in a gene g transcribed and thus expressed (as RNA) in a cell?

One solution: Measure RNA levels (result of transcription)

Method: Microarray or RNA Seq array

Result: for each subject tissue sample s, obtain a feature vector:

$$\Phi(s) = \mathbf{x} = (x_1, \ldots, x_{20,000})$$

consisting of expression levels of 20,000 genes.

Can we classify tissues this way?

Goals:

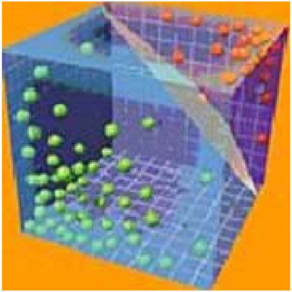
1. Differentiate two different but similar cancers.
2. Understand genetic pathways of cancer

Basic difficulties: few samples (e.g., 30-200); high dimension (e.g., 5,000 - 100,000).

Curse of dimensionality - too few samples and too many parameters (dimensions) to fit them.

Tool: Support vector machine (SVM)

Procedure: look at feature space F in which $\Phi(s)$ lives, and differentiate examples of one and the other cancer with a hyperplane:



Methods needed for full analysis (of SVM and other high dimensional methods):

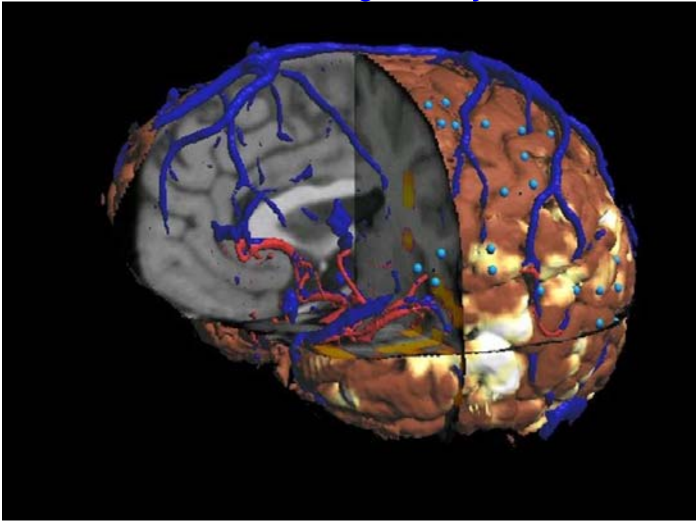
Reproducing kernel Hilbert spaces (RKHS)

2. Machine Learning: The role of *learning theory*

The role of learning theory has grown a great deal in:

- Mathematics
- Statistics
- Computational Biology
- Neurosciences, e.g., theory of plasticity, workings of visual cortex

Learning theory



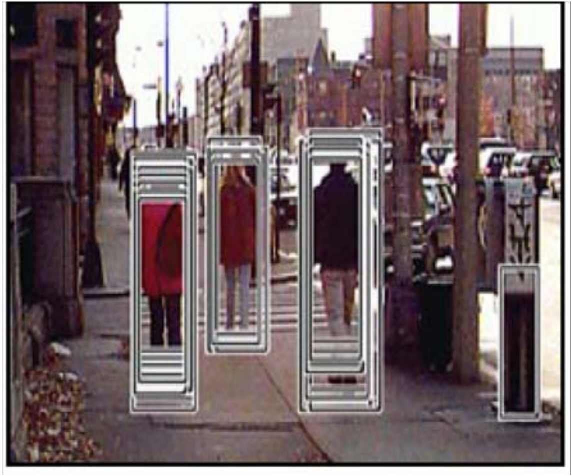
Source: University of Washington

Kernel methods

Kernel methods are used widely in:

- Computer science, e.g., vision theory, graphics, speech synthesis

Kernel methods



Source: T. Poggio/M

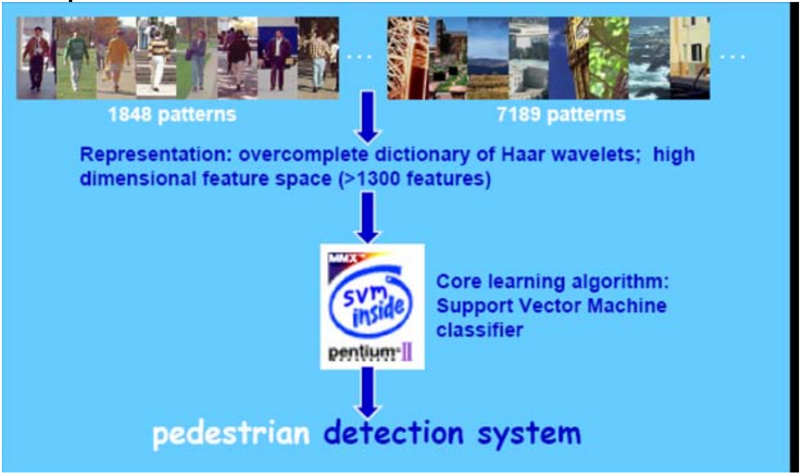
Kernel methods

Face identification:

MIT

Kernel methods

People classification or detection:



Poggio/MIT

Learning theory

We want the theory behind such learning algorithms-

3. The learning theory problem

Given an unknown function $f(\mathbf{x}) : \mathbb{R}^d \to \mathbb{R},$ learn $f(\mathbf{x})$ from a few examples, i.e., a few inputs **x** where $f(\mathbf{x})$ is known.

Determine unknown $f(\mathbf{x})$ from knowing its value at several points **x**.

Learning theory

Example 1: $\mathbf{x}$ is retinal activation pattern (i.e., $x_i =$ activation level of retinal neuron i), and $y = f(\mathbf{x}) > 0$ if the retinal pattern is a chair; $y = f(\mathbf{x}) < 0$ otherwise.

[Thus: want concept of a chair]

Given: examples of chairs (and non-chairs): $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, together with proper outputs $y_1, \ldots, y_n$. The information is in a training set $\mathcal{T} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$

This is the information:

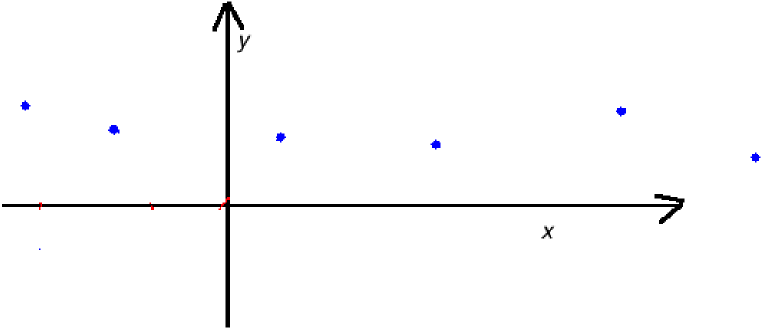
$$Nf = (f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n))$$

Learning theory

Goal: Give best possible estimate of the unknown function f, i.e., try to learn the concept f from the examples in $\mathcal{T}$. $N f$.

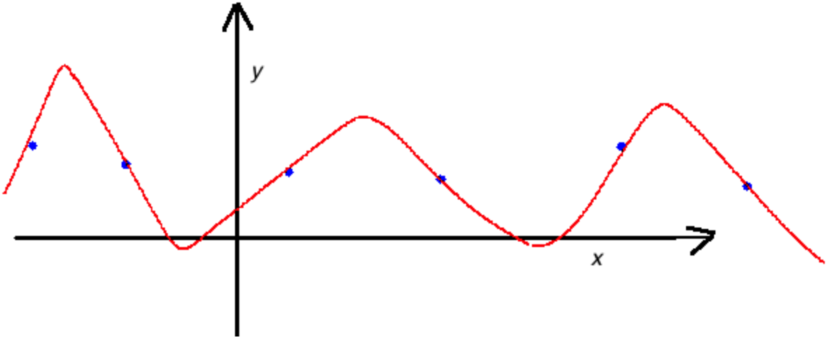
But: given pointwise information about f not sufficient: which is the "right" $f(x)$ given the training data points $\mathcal{T}$ $N f$ below?

Learning theory



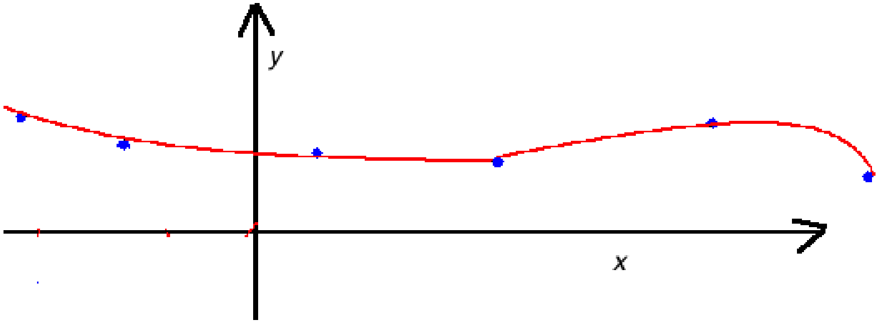
Learning theory

(a)



Learning theory

(b)



[How to decide?]

4. Infinite dimensional vector spaces:

[This material is short course in real/functional analysis; see me if you want more sources]

[Notation: in infinite dimensions generally don't use boldface on vectors]

Let H be a vector space with inner product. Recall by definition

$$\langle v, v \rangle = \|v\|^2.$$

Infinite dimensional spaces

Recall $\|v\| =$ norm $v =$ length v.

Distance between vectors v_1, v_2: $\|v_1 - v_2\|$.

Consider infinite collection

$$S = \{v_1, v_2, v_3 \ldots\} \subset H.$$

Infinite dimensional spaces

Define infinite linear combinations by:

$$\sum_{i=1}^{\infty} c_i v_i = w$$

if

$$\left\| w - \sum_{i=1}^{n} c_i v_i \right\| \xrightarrow[n \to \infty]{} 0.$$

[Definitions of span, linear independence, basis same except we now allow infinite sums]

Infinite dimensional spaces

Def. 4. All previous linear algebra definitions (e.g. spanning, linear independence, basis) extend directly to the case of infinite numbers of vectors.

Example: A collection $\{v_1, v_2, \dots\}$ of vectors *spans* a vector space V if every vector $v \in V$ can be written as a (possibly infinite) linear combination $v = c_1 v_1 + c_2 v_2 + \dots = \sum_{i=1}^{\infty} c_i v_i$.

[Henceforth always allow infinite linear combinations.]

Hilbert spaces

Def 5: An inner product space H is *complete* if any sequence $\{x_i\}_{i=1}^{\infty} \subset H$ which is *Cauchy*, i.e., $\|x_i - x_j\| \underset{i,j\to\infty}{\longrightarrow} 0$ (that is, it *should* converge) actually converges to some $x \in H$, i.e.

$$x_i \longrightarrow x.$$

[Thus if the sequence bunches up, there is something for it to converge to.]

Such an inner product space H that is complete is called a Hilbert space.

Example of incomplete space

Ex: Not all inner product spaces are Hilbert spaces since not all are complete. As an example, consider the space $P = \{$all polynomials on $[0,1]\}$. Define inner product $(f,g) = \int_0^1 f(x)g(x)dx$.

Then resulting norm is $||f(x)||^2 = \langle f, f \rangle = \int_0^1 f^2(x)dx$.

This space is not complete: consider the vectors v_N defined by the partial sums of the Taylor series for e^x :

Example of incomplete space

$$v_N = \sum_{n=0}^{N} \frac{x^n}{n!}$$ = a polynomial.

Example of incomplete space

Note that if $N > M$ then

$$||v_N - v_M|| = \left\| \sum_{n=0}^{N} \frac{x^n}{n!} - \sum_{n=0}^{M} \frac{x^n}{n!} \right\| = \left\| \sum_{n=M+1}^{N} \frac{x^n}{n!} \right\| \leq \sum_{n=M+1}^{N} \left\| \frac{x^n}{n!} \right\|$$

But:

$$\left\| \frac{x^n}{n!} \right\| = \frac{1}{n!} \|x^n\| = \frac{1}{n!} \left(\int_0^1 x^{2n} dx \right)^{1/2} = \frac{1}{n!} \left(\frac{1}{2n+1} \right)^{1/2}.$$

Example of incomplete space

So easy to show $\sum_{n=0}^{\infty} \| \frac{x^n}{n!} \| < \infty$. Thus it easily follows that

$$||v_N - v_M|| \xrightarrow[N, M \to \infty]{} 0,$$

so that the sequence v_N is a Cauchy sequence in H.

Example of incomplete space

But note that by Taylor series

$$v_N(x) - e^x = \sum_{n=0}^{N} \frac{x^n}{n!} - e^x \underset{N \to \infty}{\longrightarrow} 0$$

uniformly on [0,1]. Thus easy to show that

$$\|v_N(x) - e^x\| \underset{N \to \infty}{\longrightarrow} 0.$$

Example of incomplete space

So:

$$v_N(x) \to e^x.$$

But: can show that a sequence of functions can't converge to 2 different functions. Thus there is no polynomial $p(x)$ (i.e. something in our space P) such that

$$v_N(x) \to p(x).$$

Thus v_N do not converge to something in P and thus P is *not* complete!

[Moral: intuitively, complete space is one where any convergent sequence P_n converges to an element P *of the original space*.]

Theorem 4: If $B = \{v_1, v_2, v_3, \dots\}$ is a collection of vectors that is orthonormal (i.e, unit lengths and inner product 0), then it is automatically linearly independent.

If B is a basis for H and is orthonormal, it is called an *orthonormal basis*.

Examples of Hilbert spaces

Ex 2: $H = \mathbb{R}^3 = \{v = (v_1, v_{2,} v_3) | v_i \in \mathbb{R}\}$ is a Hilbert space (i.e., not hard to show that it's complete). Inner product is the usual one for vectors:

$$(v, w) = v_1 w_1 + v_2 w_2 + v_3 w_3.$$

This H is a Hilbert space.
Orthonormal basis:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Examples of Hilbert spaces

Ex 3:

$H = \mathbb{P}^2 =$ second order polynomials on $[0,1] =$

$$\{a_0 + a_1 x + a_3 x^2 : a_i \in \mathbb{R}\}$$

forms a Hilbert space.

Inner product:

$$(p_1(x), p_2(x)) = \int_0^1 p_1(x) p_2(x)\, dx.$$

Examples of Hilbert spaces

Note it is not hard to show that H is complete (in fact any finite dimensional vector space is complete).

Thus H is a Hilbert space.

Ex 4: Note $H = \mathbb{R}^\infty = \{v = (v_1, v_2, v_3, \dots)|v_i \in \mathbb{R}\}$ is (almost) a Hilbert space, if we define the inner product

$$(v, w) = v_1 w_1 + v_2 w_2 + \dots = \sum_{i=1}^{\infty} v_i w_i$$

Examples of Hilbert spaces

Length of a vector v is

$$\|v\| = \sqrt{\sum_{i=1}^{\infty} v_i v_i} = \sqrt{\sum_{i=1}^{\infty} v_i^2} \,.$$

Thus to have well-defined lengths we add to the definition of H, the condition that

$$\|v\| < \infty$$

for all $v \in H$. Then can show that H satisfies all the properties of a Hilbert space (in particular it's complete).

Examples of Hilbert spaces

Can show that the set of vectors

$$
\begin{aligned}
v_1 &= (1, 0, 0, \dots) \\
v_2 &= (0, 1, 0, \dots) \\
v_3 &= (0, 0, 1, \dots) \\
&\vdots
\end{aligned}
$$

is certainly orthonormal, and it spans H, so it is an orthonormal basis for H.