Learning functions on graphs and manifolds; Application to Psychological testing

Edo Libetry

September 21, 2006

Introduction

Fourier series as a 1d kernel method
- Fourier series as a 1d kernel method
- The graph Laplacian kernel for functions over vertices of undirected graphs
- The diffusion kernel for functions over manifolds

Psychological testing
- Common psychological tests
- MMPI-2 Structure and scoring

Applying the kernel technique to the MMPI-2 test
- Assumptions and motivations
- Application details

Experimental results

Discussion

Future work

Introduction

Fourier series as a 1d kernel method
- Fourier series as a 1d kernel method
- The graph Laplacian kernel for functions over vertices of undirected graphs
- The diffusion kernel for functions over manifolds

Psychological testing
- Common psychological tests
- MMPI-2 Structure and scoring

Applying the kernel technique to the MMPI-2 test
- Assumptions and motivations
- Application details

Experimental results

Discussion

Future work

Introduction

- We view a psychological evaluation of a number of people as a noisy function of their response to a test.
- We denoise the diagnosis function using the diffusion/heat kernel.
- We then extend our denoised version to the rest of the space.
- Finally the learned diagnostic function is used to score new cases.

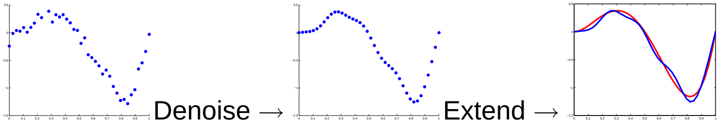


Figure: A simplified sketch of the learning by extension process.

Introduction

Fourier series as a 1d kernel method

- Fourier series as a 1d kernel method
- The graph Laplacian kernel for functions over vertices of undirected graphs
- The diffusion kernel for functions over manifolds

Psychological testing

- Common psychological tests
- MMPI-2 Structure and scoring

Applying the kernel technique to the MMPI-2 test

- Assumptions and motivations
- Application details

Experimental results

Discussion

Future work

Fourier series as a 1d kernel method

- We are given the values of a function $f : [0,1] \to \mathbb{R}$ on n points from $[0,1]$, $x_i = \frac{i-1}{n-1}$, $n \in 1 \ldots n$.
- We are also told that the values have been added random noise.
- We want to find an approximation to f, namely $\widetilde{f}$.
- Such that $\|f - \widetilde{f}\|$ is sufficiently small.
- And that $\widetilde{f}$ is somewhat smooth as well.

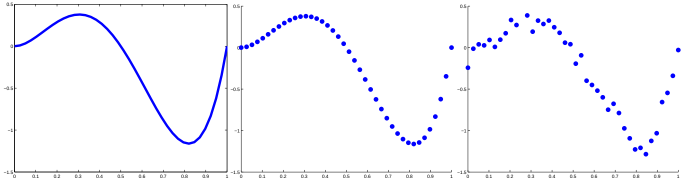


Figure: Left: original function. Center: sampled function. Right: noise added.

Fourier series as a 1d kernel method

- Decide on a measure of smoothness.
- For example minimizing the discrete curvature $c = \sum(-f(x_{i+1}) + 2f(x_i) - f(x_{i-1}))^2$.
- $c = fLf^T$ where L is a symmetric PSD matrix.
- The functions that minimize this operator are the eigenfunction of L.
- $u_k(x_i) = \sqrt{\frac{2}{n-1}} sin(k\pi x_i)$
- Constructing an approximation to f as a linear combination of these functions is easy $f \approx \widetilde{f} = \sum_{i=1}^{m} \langle u_k, f \rangle u_k$. (due to the orthogonality of the u_ks.)
- If m is small then the curvature of $\widetilde{f}$ is also bounded.

Fourier series as a 1d kernel method

1. Define a relevant operator, K ,on the sampled points.
2. Calculate the eigenfunctions (or eigenvectors) U of K.
3. Expand f as a linear combination of a small subset of U.
   $\widetilde{f}(x_i) = \sum_{k=1}^{m} \langle u_k, f \rangle u_k(x_i)$
4. Using the extended eigenvectors evaluate the function on the whole space.
   $\widetilde{f}(x) = \sum_{k=1}^{m} \langle u_k, f \rangle u_k(x)$.

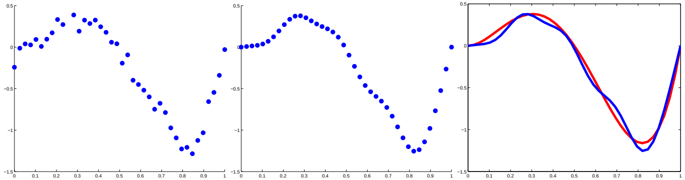


Figure: Left: The noisy sample. Center: denoised sample. Right: extension to $[0, 1]$.

The graph Laplacian kernel for functions over vertices of undirected graphs

- Given a an undirected graph $G(V,E)$ and edge weights $W_{i,j}$
- minimize the difference between neighbors with heavy weights between them. $min \sum_{(i,j)\in E} W_{i,j}(f_i - f_j)^2$.
- Define $D_{i,i} = \sum_j W_{i,j}$, and $L = D - W$ is the Graph Laplacian.
- $\sum_{(i,j)\in E} W_{i,j}(f_i - f_j)^2 = fLf^T$
- It might be appropriate to expand f as a combination of the eigenvectors of L corresponding to low eigenvalues.

The diffusion kernel for functions over manifolds

- Here we are not given a graph but a set of points X sampled from a manifold M.
- The idea is to view the values of f as the "temperatures" of the vertices. Then let them cool down or warm up according to their neighborhoods.
- This amounts to setting the kernel $K_{i,j} = e^{-\frac{\|x_i - x_j\|^2}{\sigma^2}}$
- σ here is proportional to the "time waited", or it can be viewed as a scaling parameter.
- There are a few different normalizations of K. (diffusion maps, laplacian eigenmaps).

Extending the diffusion kernel

- We now try to extend the eigenvectors of K to include x using the Nyström extension.
- since u_k are eigenvectors of K. $\lambda_k u_k(x_i) = \sum_{j=1}^{n} K(x_i, x_j) u_k(x_j)$.
- $u_k(x) = \frac{1}{\lambda_k} \sum_{j=1}^{n} K(x, x_j) u_k(x_j)$.
- Since we can evaluate $K(x, x_k)$ we can approximate $u_k(x)$.
- Notice that as $\lambda_k \to 0$ the operation becomes numerically unstable.
- We get $\widetilde{f}(x) = \sum_{k=1}^{m} \langle u_k, f \rangle u_k(x)$.

Examples

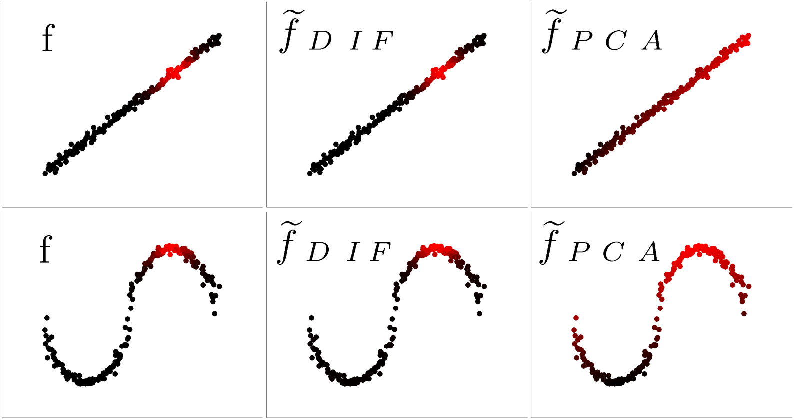


Figure: A function f approximated with the heat diffusion kernel $\widetilde{f}_{DIF}$ and the PCA kernel $\widetilde{f}_{PCA}$

Introduction

Fourier series as a 1d kernel method
- Fourier series as a 1d kernel method
- The graph Laplacian kernel for functions over vertices of undirected graphs
- The diffusion kernel for functions over manifolds

Psychological testing
- Common psychological tests
- MMPI-2 Structure and scoring

Applying the kernel technique to the MMPI-2 test
- Assumptions and motivations
- Application details

Experimental results

Discussion

Future work

Common psychological tests

- Aim to evaluate the amount of which a person possesses a certain trait, quality, or ability
  - IQ tests
  - GRE, GMAT, SAT etc.
  - Job placement tests
- Aim to evaluate the psychological state a person is in.
  - Rorschach
  - Anxiety tests, Stress Tests, etc.
- Aim to diagnose a Psychological pathology, or disorder
  - MMPI-2TM (Minnesota Multiphasic Personality Inventory-2TM)
  - BPS (Bricklin Perceptual Scales)
  - CPI (California Personality Inventory)
  - .... and many more.
- Personality tests
  - Myers-Briggs, MMPI-2, BPS, colorTest, etc.

MMPI-2 Structure and scoring

- 567 yes/no questions.
- A scale contains a subset of these questions (items) and an indicated answer (response). Usually about 40-60.
- Example: a scale measuring Depression might include the items
  - I find it hard to wake up in the morning (yes).
  - Most of the people I know are less fortunate then I'm (no).
- Each scale aims to measure one trait or condition.
- If a person answers in the indicated (keyed) way, his/her score on that scale is incremented.
- T scores. Are the number of items answered in the keyed way.
- The scales were constructed according to linear correlation between items and conditions.

Problems with the scoring method

- No reason all items should be weighted equally.
- No reason the function from answer vector to score should be linear.
- The method is highly sensitive to missing answers.
- The correlation between item and scale is sometime debatable.
- New scales and changes to the existing ones are very frequent.

MMPI-2 score interpretation

- T scores are normalized (mean 50, std 10).
- Only scores above 65, are considered abnormal.
- Scores above 60 are considered ”Elevated”.
- The set of elevated scales for a person is called his/her ”Type” or ”Profile”.
- Certain types are well known and studied and others are ambiguous (at best).

Introduction

Fourier series as a 1d kernel method
- Fourier series as a 1d kernel method
- The graph Laplacian kernel for functions over vertices of undirected graphs
- The diffusion kernel for functions over manifolds

Psychological testing
- Common psychological tests
- MMPI-2 Structure and scoring

Applying the kernel technique to the MMPI-2 test
- Assumptions and motivations
- Application details

Experimental results

Discussion

Future work

Assumptions and motivations

- Assumptions:
  - We assumed that similar types give similar responses.
  - The set of responses lay on a low dimensional manifold, with additive noise.
  - The MMPI-2 scores and types are smooth on that manifold.
- Motivations:
  - Test the validity and utility of our assumptions.
  - Find a natural set of functions that spans (and denoises) the space of psychological diagnosis.
  - Find the types and disorders as clusters or areas on the manifold.
  - Hopefully devise an alternative scoring method.

Application details

- The diffusion kernel was chosen.
- The distance between responses was taken to be the Hamming distance.
- The MMPI-2 score functions were learned using a training set of 500 subject.
- The functions were evaluates on a test set of a 1000 other subjects.
- The score approximation was done using 15 eigenvectors.
- Correlations between f and $\widetilde{f}$ were calculated.
- Missing answers 1: The same experiment was repeated while randomly deleting answers to check the robustness of our method.
- Missing answers 2: Each scale was predicted after deleting all the items that correspond to it.

Introduction

Fourier series as a 1d kernel method
- Fourier series as a 1d kernel method
- The graph Laplacian kernel for functions over vertices of undirected graphs
- The diffusion kernel for functions over manifolds

Psychological testing
- Common psychological tests
- MMPI-2 Structure and scoring

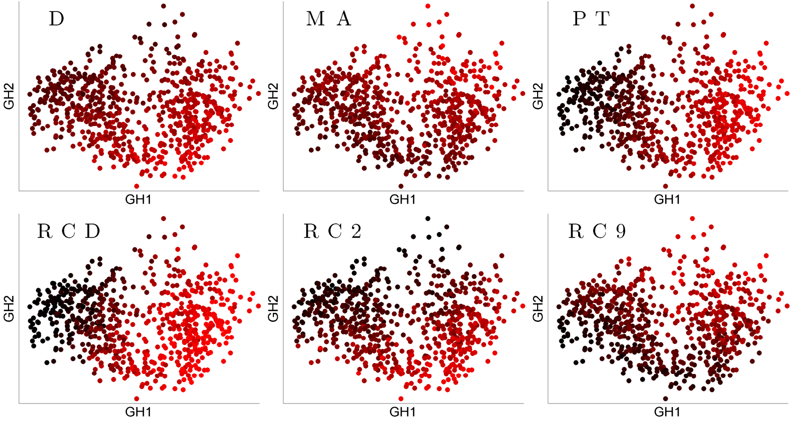
Applying the kernel technique to the MMPI-2 test
- Assumptions and motivations
- Application details

Experimental results

Discussion

Future work

Scores over the diffusion map



Number of eigenvectors

- Using the first 15 eigenvectors of the diffusion kernel one can achieve a good approximation of the MMPI-2 scale scores.

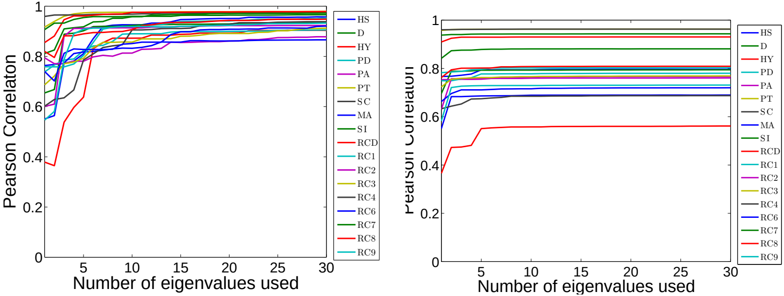


Figure: Pearson Correlations between f and $\widetilde{f}$ for different numbers of eigenvalues used. For comparison, on the righthand side, the same plot using the PCA kernel.

Missing data results

	q=30		q=50		q=100		q=200		q=300	
	r	Hit rate	*r*	Hit rate	*r*	Hit rate	*r*	Hit rate	*r*	Hit rate
HS	95	89	95	89	94	87	94	83	92	80
D	93	83	93	83	93	83	92	80	92	77
HY	89	71	88	70	88	69	87	67	84	62
PD	91	76	91	77	91	76	90	74	89	71
PA	88	67	88	67	87	66	87	64	85	62
PT	98	97	98	97	98	97	98	97	97	96
SC	98	98	98	98	98	98	98	98	97	97
MA	87	67	87	67	86	67	86	64	85	65
SI	96	91	96	92	96	91	95	90	95	88
RCD	98	96	97	96	97	96	97	94	97	94
RC1	95	86	94	86	94	85	93	81	91	75
RC2	93	82	93	82	92	80	92	79	91	77
RC3	89	73	89	73	89	72	89	71	88	70
RC4	92	81	92	78	92	76	90	74	88	69
RC6	92	78	92	78	91	78	91	75	89	72
RC7	96	92	96	93	96	92	96	91	95	91
RC8	93	84	93	84	93	83	93	82	91	78
RC9	93	82	93	81	92	80	92	77	91	75

Table: Pearson correlations, r, between f and $\widetilde{f}$. Here q items were randomly deleted from each answers sequence. The hit rate indicated is the percent of subjects classified within $1/2$ standard deviation from their original score. The variance of the correlations and hit rates, for different choices of the base data set, is smaller the 0.02.

Missing scale results

- In the table below each scale score was estimated while all the items that belong to that scale were missing.
- For comparison we tried also to complete the missing responses using a Markov process and score the corrupted records using the usual scoring procedure.

Scale	r_{GH}	Hit $rate_{GH}$	r_{MC}	Hit $rate_{MC}$	Scale	r_{GH}	Hit $rate_{GH}$	r_{MC}	Hit $rate_{MC}$
HS	79	59	69	46	RCD	94	85	60	35
D	86	67	65	0	RC1	77	57	74	57
HY	74	51	55	0	RC2	87	67	49	34
PD	80	59	48	0	RC3	81	59	39	36
PA	78	54	55	5	RC4	67	48	30	34
PT	94	88	70	26	RC6	79	62	41	38
SC	94	85	73	41	RC7	92	81	60	40
MA	80	58	35	2	RC8	87	70	56	47
SI	87	69	58	7	RC9	86	67	32	26

Table: The variance of the correlations and hit rates, for different choices of the base data set, is smaller the 0.02.

Introduction

Fourier series as a 1d kernel method
- Fourier series as a 1d kernel method
- The graph Laplacian kernel for functions over vertices of undirected graphs
- The diffusion kernel for functions over manifolds

Psychological testing
- Common psychological tests
- MMPI-2 Structure and scoring

Applying the kernel technique to the MMPI-2 test
- Assumptions and motivations
- Application details

Experimental results

Discussion

Future work

Discussion

- Using the above method proved to be a useful tool in scoring personality tests.
- It has proven itself superior to other methods when dealing with corrupted responses.
- It has not been shown in the results section but one can score profiles directly from the test without the use of scales (a task unachieved by psychologists).
- It allows to administer a shorter test while not losing diagnostic ability. Saving time and money for clinicians and institutes alike.
- We have not found any well defined clusters in responses. This contradicts theories of ”personality categorization” (Myers-Briggs, astrology, etc.)

Introduction

Fourier series as a 1d kernel method
- Fourier series as a 1d kernel method
- The graph Laplacian kernel for functions over vertices of undirected graphs
- The diffusion kernel for functions over manifolds

Psychological testing
- Common psychological tests
- MMPI-2 Structure and scoring

Applying the kernel technique to the MMPI-2 test
- Assumptions and motivations
- Application details

Experimental results

Discussion

Future work

Future work

- We claimed that 15 eigenvectors were a suitable choice. Yet, this was supported by experimental justification only.
- The assumption is that the difference between f and $\widetilde{f}$ corresponds to the noise in the function f.
- One can automate this process if it could be checked that the remainder is indeed noise.
- I now focus on eigenvalue concentration results for random matrices and other randomity tests under different models of independence.