Basis and Spectral Representations

School of Computer Science
Tel-Aviv University

October 25, 2012

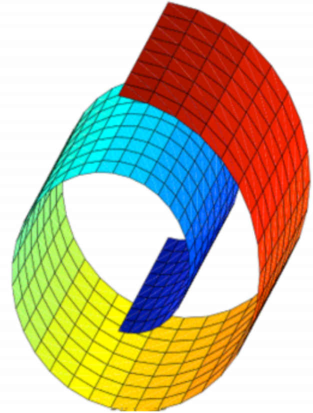
Introduction

The course will focus on modern and advanced methods in scientific computing with applications to mathematics, computer science and engineering. Lectures will be given on the following topics:

- Sparse Representations
- Compressed Sensing
- Fast Multi-pole method & Fast Gauss Transform
- Matrix Perturbation
- Matrix Completion
- Patch-Tensor Embedding
- Diffusion Maps
- Local Diffusion Folders
- Prolate Spheroidal Wave Functions

Motivation

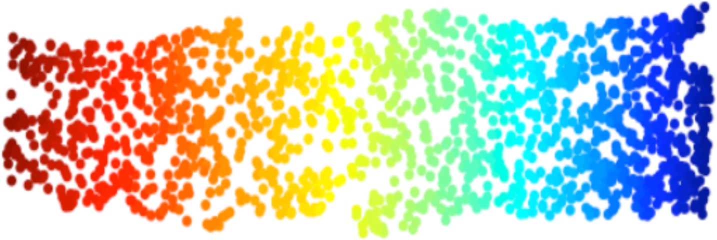
Manifold learning:The goal



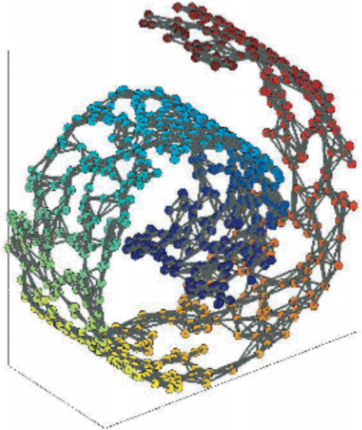
Motivation

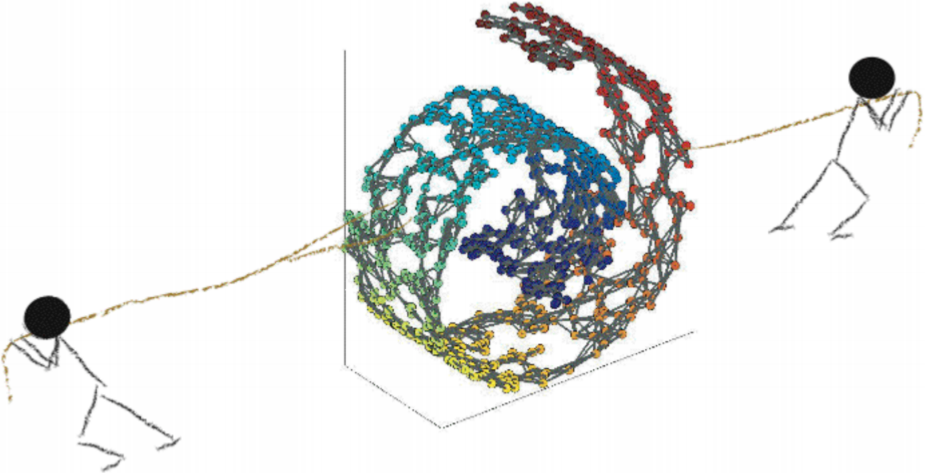
Manifold learning:The goal

Motivation
Manifold learning:The goal



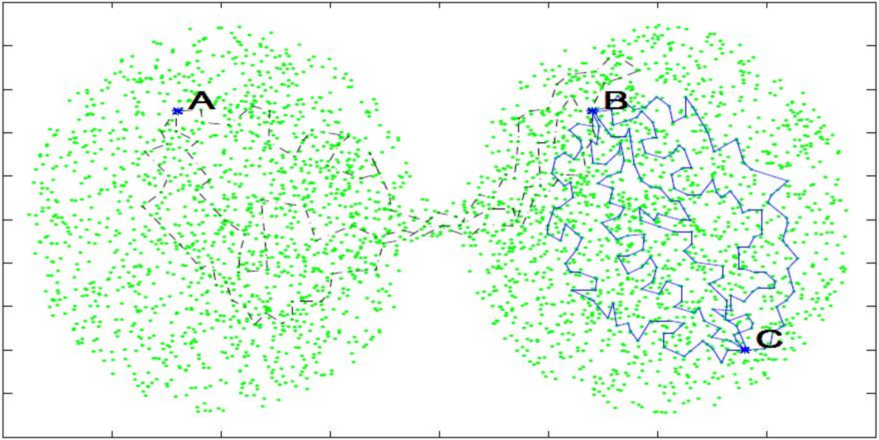
Motivation
Manifold learning:Kernel methods





Motivation

Manifold learning:Diffusion maps



Matrix Approximation and Completion

- Matrix completion is a popular research field with many applications to biology, machine learning, image processing and other scientific fields.
- In particular - Low Rank completions
- Example: The Netflix problem for movie recommendations:
  - Rows correspond to viewers, columns to movies
  - Entry $X_{ij} \in \{1, ..., 5\}$ is the rating
  - $480,000$ viewers $\times 18,000$ movies $\Rightarrow 8.9 \times 10^9$ entries
- Each viewer recommends 200 movies on average, so only 1.2% of the entries contain data
- The task is to predict the ratings that viewers will give to movies they have not yet rated

Matrix Approximation and Completion

Mathematically, the problem is:

$$\begin{aligned} &\text{minimize } \operatorname{rank}(\mathbf{X}) \\ &\text{subject to } X_{ij} = M_{ij}, \quad (i,j) \in \Omega \end{aligned} \tag{1}$$

We are looking for a matrix with minimal rank, such that the given entries do not change. The problem in Eq. 1 is NP-hard Complete by minizing the rank:

$$\mathbf{M} = \begin{bmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & ? \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & ? \end{bmatrix}.$$

The missing entry should be 6 and the rank will be 1. In this case, the solution is unique.

Matrix Approximation and Completion

In the course we will discuss some matrix completion related topics:

- Least squares approximation of matrices under different spectral regularizations
- Methods for approximating just some of the entries
- Discuss convergence of different methods: local and global convergence
- Discuss today's state of the art methods for matrix completion and their drawbacks

Fast Multi-pole Method

The FMM is a method developed to speed up computations on n-body problem.

- Developed by Greengard and Rokhlin in the late 80s
- Considered by some one of the top ten most important algorithms of the 20th century
- Can be used to accelerate matrix-vector multiplication with certain structure from $\mathcal{O}(n^2)$ to $\mathcal{O}(n)$
- Can be used to accelerate iterative methods for solving linear system of equations that perform matrix-vector multiplication (for example, conjugate gradient method).
- Also can be used for electromagnetic problems
- Can be utilized for Fast Gauss Transform

Prolate Spheroidal Wave Functions

Bandlimited functions are important in many fields of mathematics and engineering.

- PSWFs are the eigenfunctions of the bandlimited Fourier transform:

$$\int_{-1}^{1} e^{ic\,\omega\cdot x}\,\psi_n^c(\omega)\,d\omega = \lambda_n^c\,\psi_n^c(x)$$

- c is a fixed bandlimit. $x \in [-1, 1]$, $\omega \in [-1, 1]$

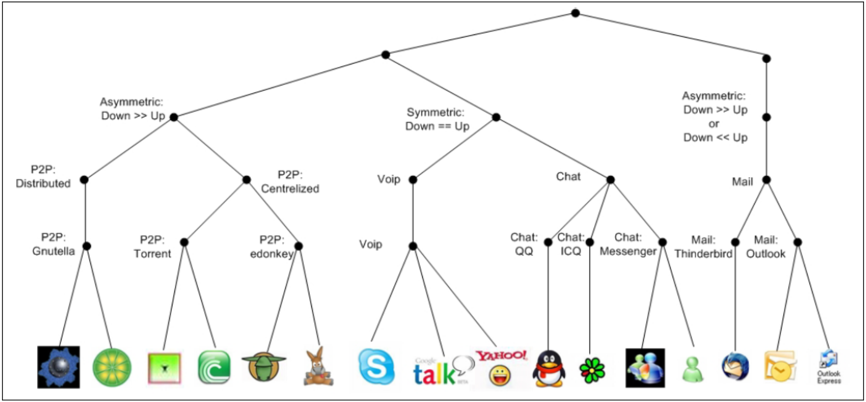
Prolate Spheroidal Wave Functions

The PSWFs developed in the 60s by David Slepian and are used in the following fields:

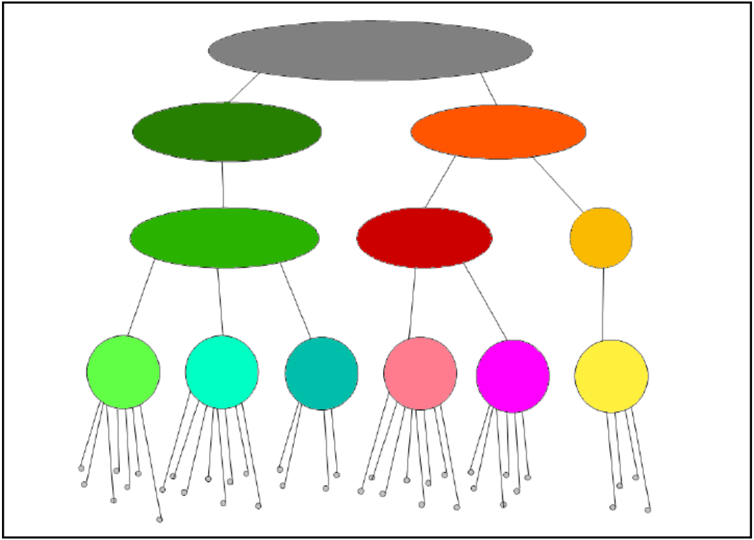
- Signal Processing: Filter design, signal reconstruction
- Interpolation
- Approximation

We will focus mainly on the properties of the PSWFs and their application to approximation of bandlimited functions (Shkolnisky & Tygert)

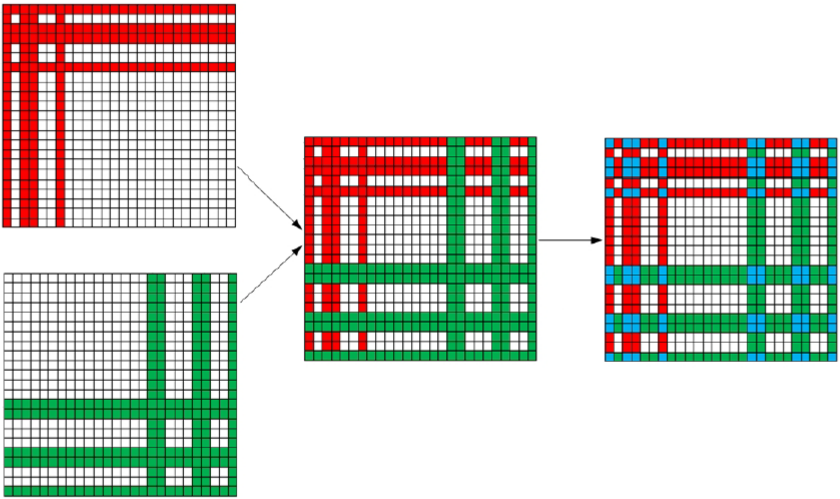
Localized diffusion folders



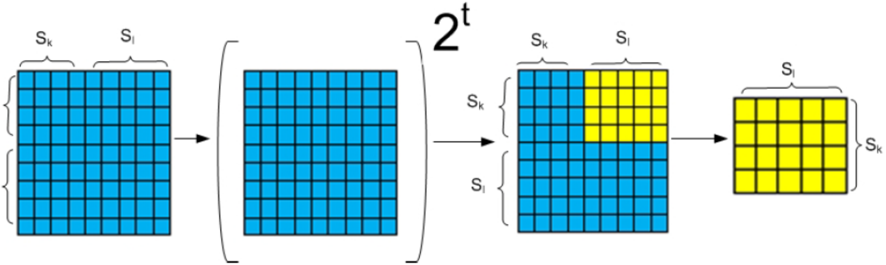
Localized diffusion folders



Localized diffusion folders

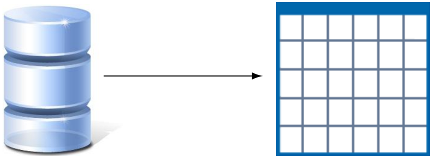


Localized diffusion folders

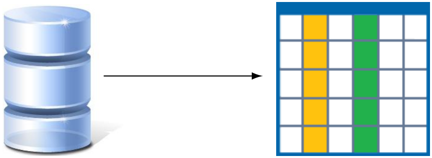


Diffusion Maps

Diffusion Maps



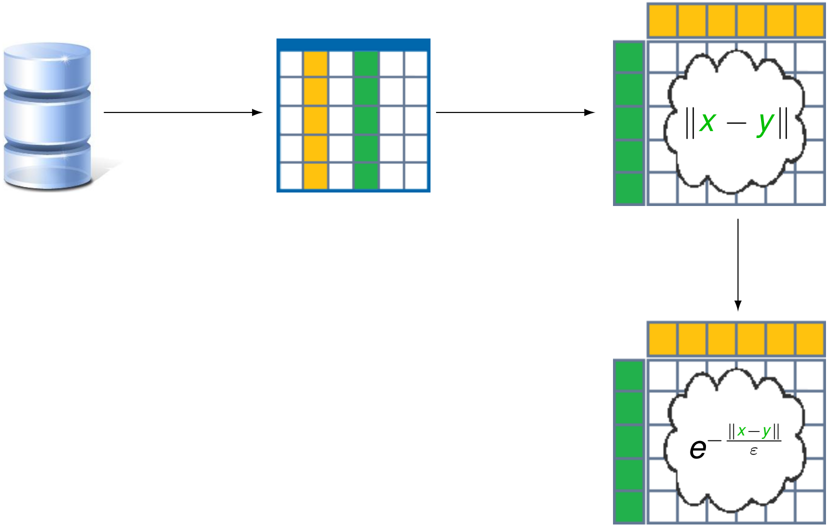
Diffusion Maps



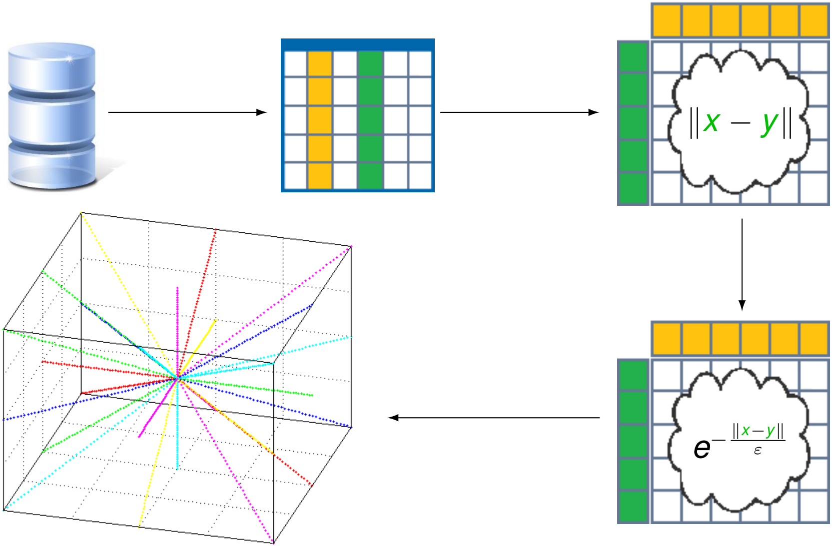
Diffusion Maps

Diffusion Maps

Diffusion Maps



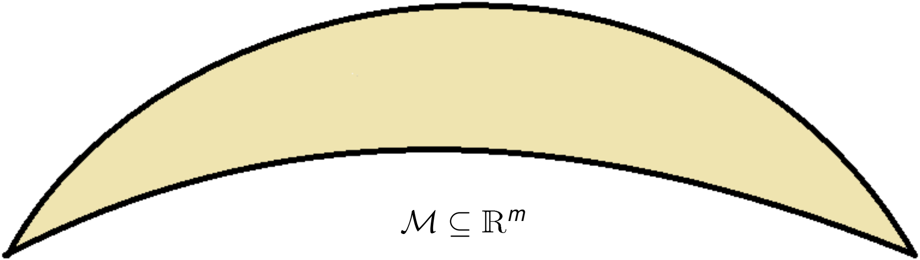
Diffusion Maps



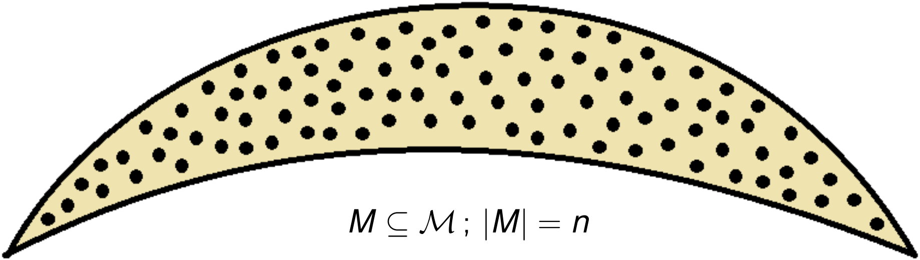
Patch to Tensor Embedding

How can we discriminate between two similar data points that have dissimilar neighborhood?

Patch to Tensor Embedding

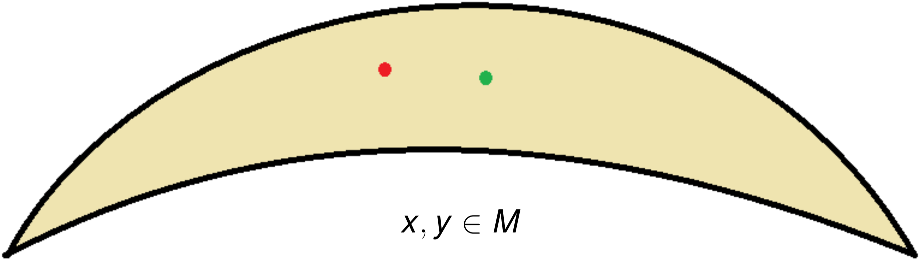


Patch to Tensor Embedding



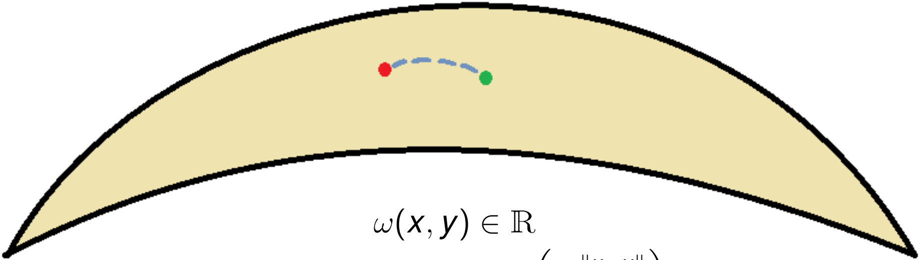
$M \subseteq \mathcal{M}\,;\ |M| = n$

Patch to Tensor Embedding



$x, y \in M$

Patch to Tensor Embedding

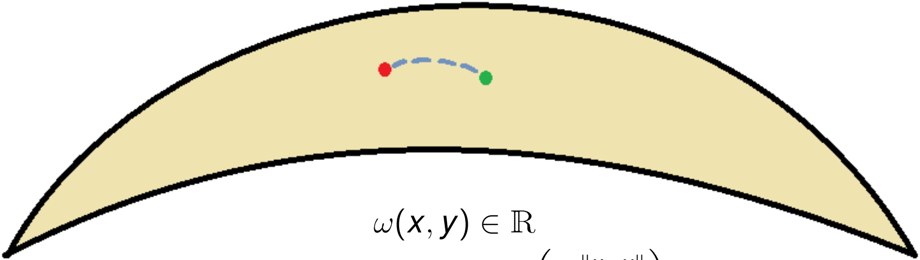


$$\omega(x, y) \in \mathbb{R}$$

$$\text{e.g., } \omega(x, y) = \exp\left(-\frac{\|x-y\|}{\varepsilon}\right)$$

Patch to Tensor Embedding

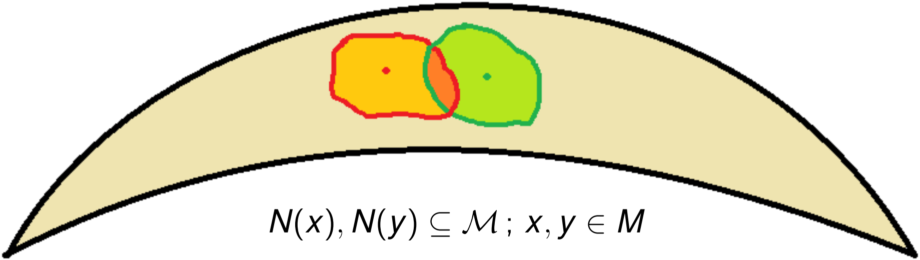
$$u \in M \mapsto \Psi(u) \in \mathbb{R}^{\ell \ll m}$$



$$\omega(x, y) \in \mathbb{R}$$

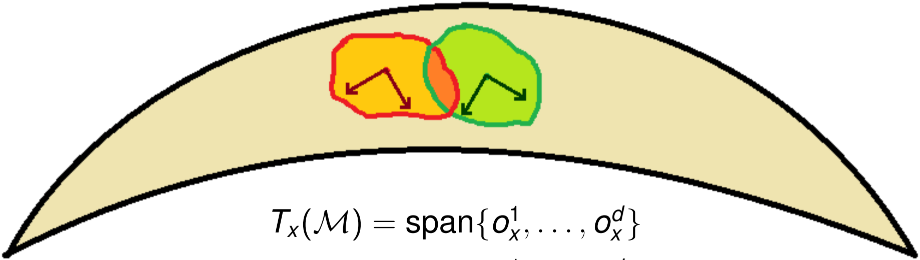
$$\text{e.g., } \omega(x, y) = \exp\left(-\frac{\|x-y\|}{\varepsilon}\right)$$

Patch to Tensor Embedding



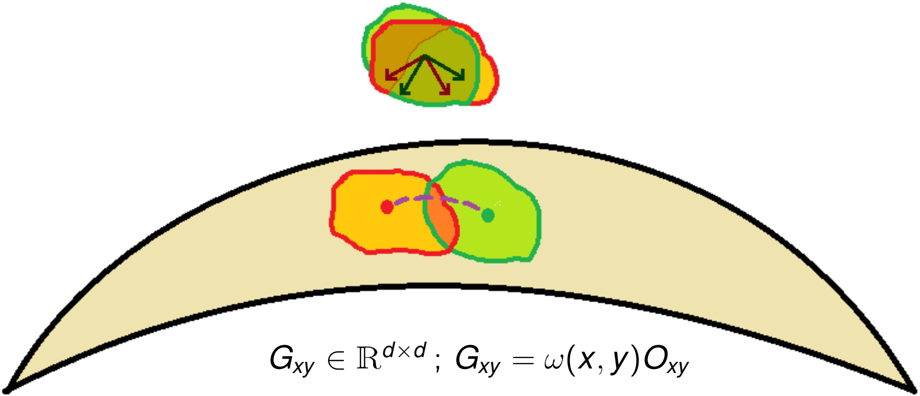
$$N(x), N(y) \subseteq \mathcal{M} \, ; \; x, y \in M$$

Patch to Tensor Embedding



$$T_x(\mathcal{M}) = \text{span}\{o_x^1, \ldots, o_x^d\}$$
$$T_y(\mathcal{M}) = \text{span}\{o_y^1, \ldots, o_y^d\}$$

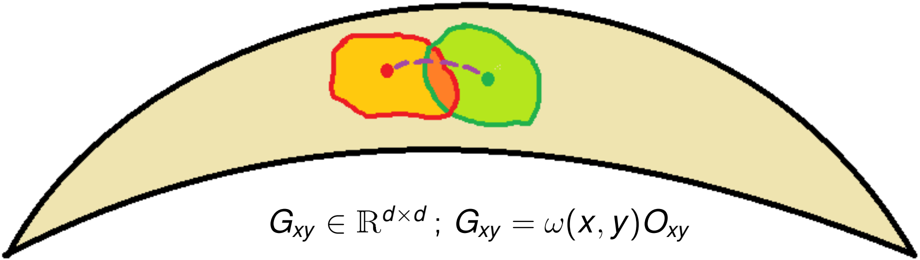
Patch to Tensor Embedding



$$G_{xy} \in \mathbb{R}^{d \times d}\,;\; G_{xy} = \omega(x, y) O_{xy}$$

Patch to Tensor Embedding

$$N(u \in M) \mapsto \mathcal{T}_u \in \mathbb{R}^{\ell} \otimes \mathbb{R}^{d}$$



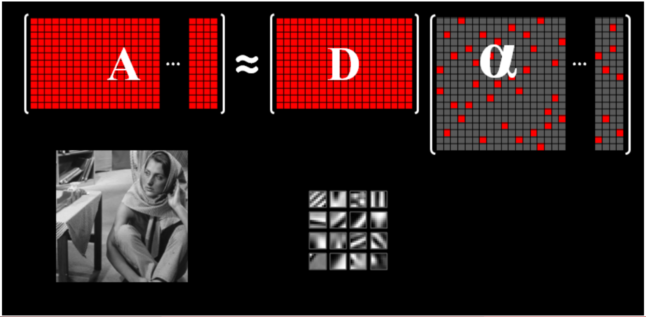
$$G_{xy} \in \mathbb{R}^{d \times d}; \; G_{xy} = \omega(x, y) O_{xy}$$

Sparse Representations and the k-SVD Algorithm

- We will overview different methods finding compact representations for matrices.That is, to present a matrix using sparse, linear combination of a small dictionary.
- Motivation : image compression, de-noising, image completion - you name it!

Sparse Representations and the k-SVD Algorithm

- In mathematical form : Given a matrix A, we would like the find matrices D and α, such that $A \approx D\alpha$ where D a column matrix of a small set of columns in A, and α is sparse.
    - Here, each column of α is a sparse linear combination of the columns of D forming a column of A.



Sparse Representations and the k-SVD Algorithm

- we will go over the k-SVD algorithm which find such representation in an elegant way, using ideas from the Singgular Value Decomposition, and k-Means algorithms.
- Based on papers and lectures of Michael Elad.

Matrix Perturbation Theory and Its Applications

- In a nutshell : how does a small change in the input affects the output?
- Given a matrix A and some function ϕ which operates on A, we are interested in understanding how a perturbation added to the matrix, affects ϕ.
- That is to understand the relation between $\phi(A)$ and $\phi(A+\epsilon)$, where ϵ is a small perturbation (can be noise).
- Interesting ϕ operators can be ones which finds the eigenvalues of A, its eigenvectors, $\|A\|$ and so on.

Matrix Perturbation Theory and Its Applications

- Example application(1) : Computation of the Google page rank algorithm - how to update the ranks without recomputing the entire algorithm.
- Example application(2) : updating a training set profile with small changes to the input set.
- Books
  - Matrix Analysis - R. Bhatia
  - Matrix Perturbation Theory - Stewart and Sun.