Sparse Coding

Xinrui Lyu

Overview

- Review: Orthogonality
- Change of Basis and Fourier transform
- Wavelet and DCT

Orthogonality

Inner product

- For $\mathbf{u}, \mathbf{v} \in \mathbb{R}^d$,

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^\top \mathbf{v} = \sum_{i=1}^{d} \mathbf{u}_i \mathbf{v}_i,$$

  is an inner product.
- Let $f(t), g(t)$ be two functions for $t \in [0, 1]$, then

$$\langle f(t), g(t) \rangle = \int_0^1 f(t) g(t) dt,$$

  is an inner product in a function space.

Orthogonality

Two vectors $\mathbf{u}, \mathbf{v} \in H$ are orthogonal if and only if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Orthogonal matrix

Basis

A basis of a vector space is a set of vectors with the following two properties:

1. It is linearly independent
2. It spans the space

Orthogonal matrix

A basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is called orthonormal if

$$\mathbf{v}_i^\top \mathbf{v}_j = \begin{cases} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \end{cases}$$

A matrix $\mathbf{A}$ with orthonormal columns is called an orthogonal matrix. The special case of $\mathbf{A}$ being an orthogonal matrix is important since the projection matrix becomes extremely simple since $\mathbf{A}^\top \mathbf{A} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix.

Pen&Paper

$$\hat{\mathbf{x}} = \sum_{k=1}^{\tilde{K}} z_k \mathbf{u}_k$$

If we consider only orthonormal bases, we can formulate the compression problem as picking from the original coefficients $z_1, \ldots, z_K$ a subset $\tilde{K}$ of them which minimize the approximation error.

Let σ be a permutation of indices $\{1, \ldots, K\}$ and $\hat{\mathbf{x}}_\sigma$ the function that uses the coefficients corresponding to the first $\tilde{K}$ indices of the permutation σ :

$$\hat{\mathbf{x}}_\sigma = \sum_{k=1}^{\tilde{K}} z_{\sigma(k)} \mathbf{u}_{\sigma(k)}$$

Pen&Paper

Question: Find the permutation σ^{min} which minimizes the L^2 approximation error $\left\|\mathbf{x} - \hat{\mathbf{x}}_\sigma\right\|_2^2$:

$$\sigma^{min} = \underset{\sigma}{\operatorname{argmin}} \left\|\mathbf{x} - \hat{\mathbf{x}}_\sigma\right\|_2^2$$

Hint: Remember that the L^2 norm can be written with the help of an inner product as

$$\left\|\mathbf{x} - \hat{\mathbf{x}}_\sigma\right\|_2^2 = \langle \mathbf{x} - \hat{\mathbf{x}}_\sigma, \mathbf{x} - \hat{\mathbf{x}}_\sigma \rangle$$

Pen&Paper

If basis $\mathbf{u}_1, \ldots, \mathbf{u}_k$ is orthonormal it means that

$$\begin{aligned} \langle \mathbf{u}_k, \mathbf{u}_l \rangle &= 0, \qquad k \neq l \\ \langle \mathbf{u}_k, \mathbf{u}_k \rangle &= 1 \end{aligned}$$

Question: What happens when $\langle \mathbf{u}_k, \mathbf{u}_k \rangle \neq 1$?

Energy Preserving Concept

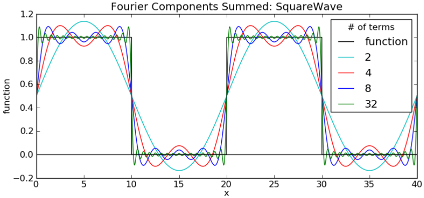
Note that the previous problem can be solved using the fact that orthonormal basis preserves the energy.

$$\forall \mathbf{x} = \mathbf{U}\mathbf{z} : \|\mathbf{x}\|_2 = \|\mathbf{z}\|_2 \tag{1}$$

Question: Does orthonormal basis preserve distances as well? What about inner product?

Why Fourier Basis for Sparse Coding

$$f(t) = \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{1}{n} \sin(\frac{k\pi t}{20})$$



DFT of a Signal

$$y = \sin 60 * 2\pi x + 1.5 \sin 80 * 2\pi x$$

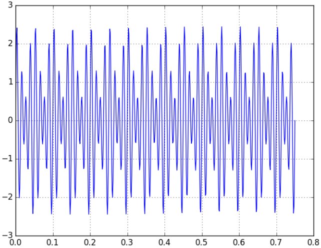


Figure: Original Signal

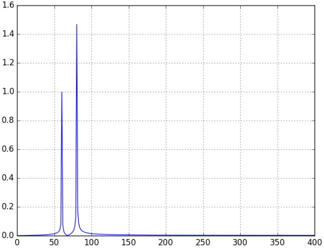


Figure: Fourier Transform

Fourier Transform of an Image

How to take FT in 2-D?

- Image can be considered as a signal in 2D
- First take FT of the columns then FT of the rows(You can interchange them)

Fourier Transform of an Image

How to take FT in 2-D?

- Image can be considered as a signal in 2D
- First take FT of the columns then FT of the rows(You can interchange them)

How to interpret FT of an image?

- Large changes in the pixel values = High frequency
- Eg : edges, background objects

FT Example

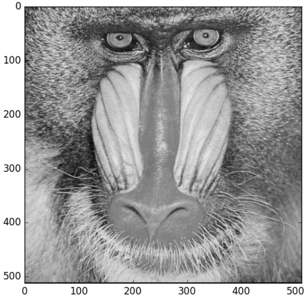


Figure: Original Image

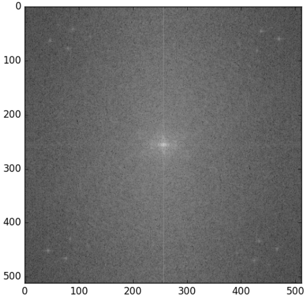


Figure: Frequency Spectrum

FT Code

```
import numpy as np
from skimage import io

#load an image of 512 by 512
im = io.imread("baboon.png")

#perform FFT
ft = np.fft.fft2(im)

#shift and compute the spectrum
shiftedft = np.fft.fftshift(ft)
spectrum = np.log10(np.abs(shiftedft))

```

Fourier Transform of an Image

- In the code, we need to shift the zero frequency component to the center of the spectrum
- Take the logarithm since the raw spectrum values are very large
- For perception, frequency components with largest magnitudes are important
- Brightness and smoothness are controlled by smaller magnitudes

Image Compression by FT

- Reconstruct image by Inverse Fourier Transform using only the frequencies with largest magnitude

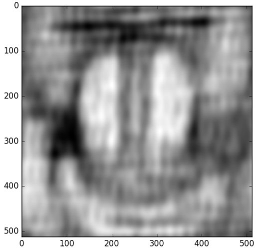


Figure: Using 0.1 percent

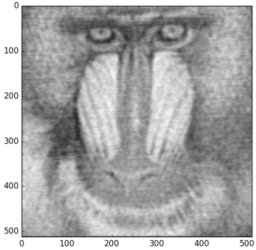
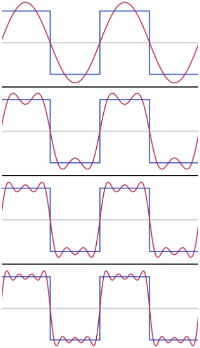


Figure: Using 1 percent

Fourier vs Wavelet

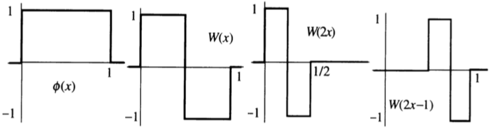
Fourier Transform



$$f(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \xi}\, dx$$

where ξ is the frequency

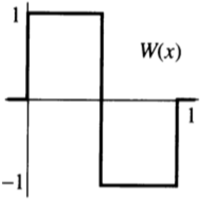
Wavelet Transform



$$X(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi\left(\frac{t-b}{a}\right) x(t)\, dt$$

where a = scaling and b = time

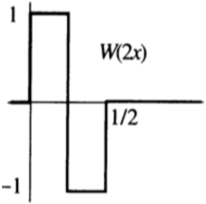
Haar Wavelets



mother wavelet

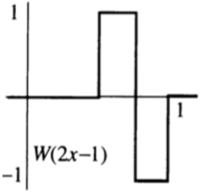
Unnormalized
Haar vector

$$\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$



$$\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$

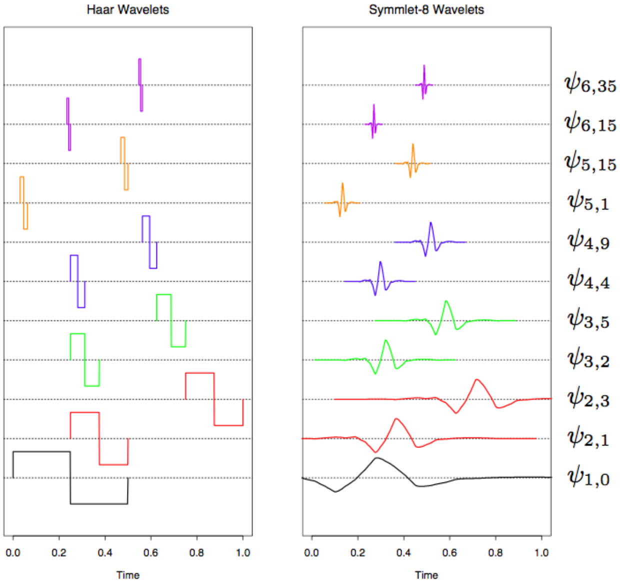
dilated



$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$$

translated

Haar Wavelets



Properties of Haar Wavelets[1]

- Mother wavelet: $\psi(t) = \begin{cases} 1 & \text{if } 0 \le t < \frac{1}{2} \\ -1 & \text{if } \frac{1}{2} \le t < 1 \\ 0 & \text{otherwise} \end{cases}$
- Haar system : $\psi_{n,k}(t) = 2^{n/2}\psi(2^n t - k), \forall n, k$ integers, $n \ge 0$
- $\psi_{n,k}(t)$ is non-zero on the interval $I_{n,k} = [k2^{-n}, (k+1)2^{-n}]$
- Has integral 0: $\int_{\mathbf{R}} \psi_{n,k}(t)dt = 0$
- Norm 1: $\|\psi_{n,k}\|^2_{L^2(\mathbf{R})} = \int_{\mathbf{R}} \psi_{n,k}(t)^2 dt = 1$
- Pairwise orthogonal: $\int_{\mathbf{R}} \psi_{n_1,k_1}(t)\psi_{n_2,k_2}(t)dt = \delta_{n_1,n_2}\delta_{k_1,k_2}$
- $\Rightarrow$ Haar system is an orthonormal basis in $L^2(\mathbf{R})$

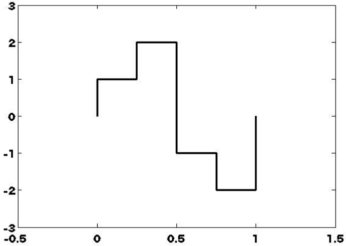
[1] `http://en.wikipedia.org/wiki/Haar_wavelet`

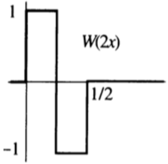
Properties of Haar Wavelets

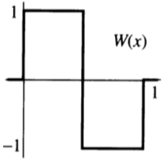
- Any continuous real function on [0, 1] can be approximated uniformly on [0, 1] by linear combinations of the constant function and Haar functions. The respective coefficients are given by: $X_\omega(k,n) = \int_{\mathbf{R}} x(t)\psi_{n,k}(t)dt$

Pen&Paper - Multiresolution Concept

Reconstruct the following signal with shifted and scaled Haar wavelets







Discrete cosine transform

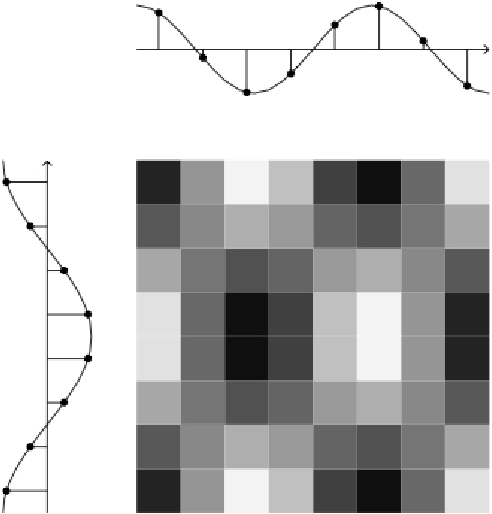
1D Discrete cosine transform:

$$z_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n+\frac{1}{2}\right)k\right] \qquad k = 0, \dots, N-1$$

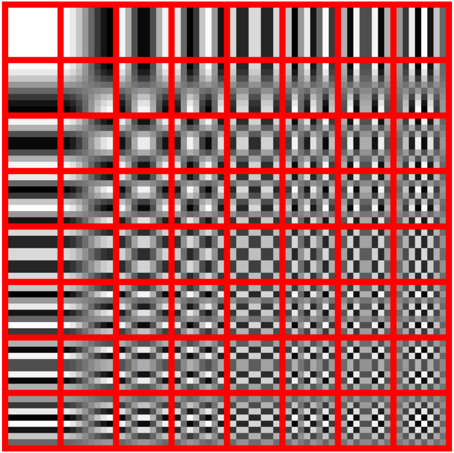
2D Discrete cosine transform:

$$\begin{aligned}
z_{k_1,k_2} &= \sum_{n_1=0}^{N_1-1}\left(\sum_{n_2=0}^{N_2-1} x_{n_1,n_2} \cos\left[\frac{\pi}{N_2}\left(n_2+\frac{1}{2}\right)k_2\right]\right)\cos\left[\frac{\pi}{N_1}\left(n_1+\frac{1}{2}\right)k_1\right] \\
&= \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} x_{n_1,n_2} \cos\left[\frac{\pi}{N_1}\left(n_1+\frac{1}{2}\right)k_1\right]\cos\left[\frac{\pi}{N_2}\left(n_2+\frac{1}{2}\right)k_2\right].
\end{aligned}$$

2-D Cosine Basis



2-D Cosine Bases



Two-dimensional DCT frequencies from the JPEG DCT