

# Natural Image Statistic: Principal Component and Whitening notes

Srikanth Gadicherla

November 26, 2016

## 1 Gaussianity as the basis for PCA

### 1.1 The Probability model related to PCA

The most interesting part of data is variances and covariances which is the central aspect of a Gaussian data, which goes by the equation

$$p(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2} |C|^{1/2}} \exp\left(-\frac{1}{2} \sum_{i,j} x_i x_j [C^{-1}]_{ij}\right) \quad (1)$$

where  $C^{-1}$  is the inverse of the covariance matrix and  $[C^{-1}]_{ij}$  is the  $i$ -th row and  $j$ -th column of that inverse matrix.

Given the data distribution is Gaussian, the covariance based methods generally suffice, but in reality the data distribution of the images is far from Gaussian.

Use PCA as statistical model given the data as Gaussian distribution, so that we could apply Bayesian inference, but for this we need to compute the prior distribution for any image data.

### 1.2 PCA as generative model

In generative models, we try to estimate the distribution(or process) which produced that data-...

### **1.3 Image Synthesis result**

Plotting the data generated from the model shows that PCA captures the general smoothness of the images. The smoothness comes from the fact that first PC corresponds to the feature which changes rather smoothly.

## **2 Power spectrum of natural images**

The covariances and the frequency based properties are related via the Wiener-Khinchin theorem

### **2.1 The $1/f$ Fourier amplitude or $1/f^2$ power spectrum**

The frequency based representation tells that power spectrum of natural images is inversely proportional to the square of the frequency.

Known that power spectrum is the square of the Fourier amplitude, so it implies that Fourier amplitude is inversely proportional to frequency.