Characteristic functions

Some theorems: $$\phi(k) = E\{e^{ikx}\} = \int_{-\infty}^{\infty} e^{ikx} p(x) dx$$

1. The characteristic function always exists
2. $$\left.\frac{d^n \phi(k)}{dk^n}\right|_{k=0} = i^n < x^n >$$
3. The characteristic function uniquely determines the pdf

$$\begin{array}{ll} \text{Binomial} & \phi(k) = \left(pe^{ik} + q\right)^n \\ \text{Poisson} & \phi(k) = e^{\lambda\left(e^{ik} - 1\right)} \\ \text{Gaussian} & \phi(k) = e^{\left(i\mu k - k^2\sigma^2/2\right)} \end{array}$$

Example: 1- Sum of two Gaussian random variables.
2- Poisson distribution with large mean approaches a Gaussian PDF.

Central limit theorem

Given a set $x_1, x_2, \ldots, x_N$ of independent, identically distributed (*iid*) random variables with zero mean and σ^2 variance then the distribution of $\sum_n x_n$ approached Gaussian with zero mean and variance σ^2 when N approaches infinity.

Proof:

$$\phi(k) = \int dx\, p(x)(1 + ikx - k^2 x^2/2 + \ldots) \approx 1 - k^2\sigma^2/2$$

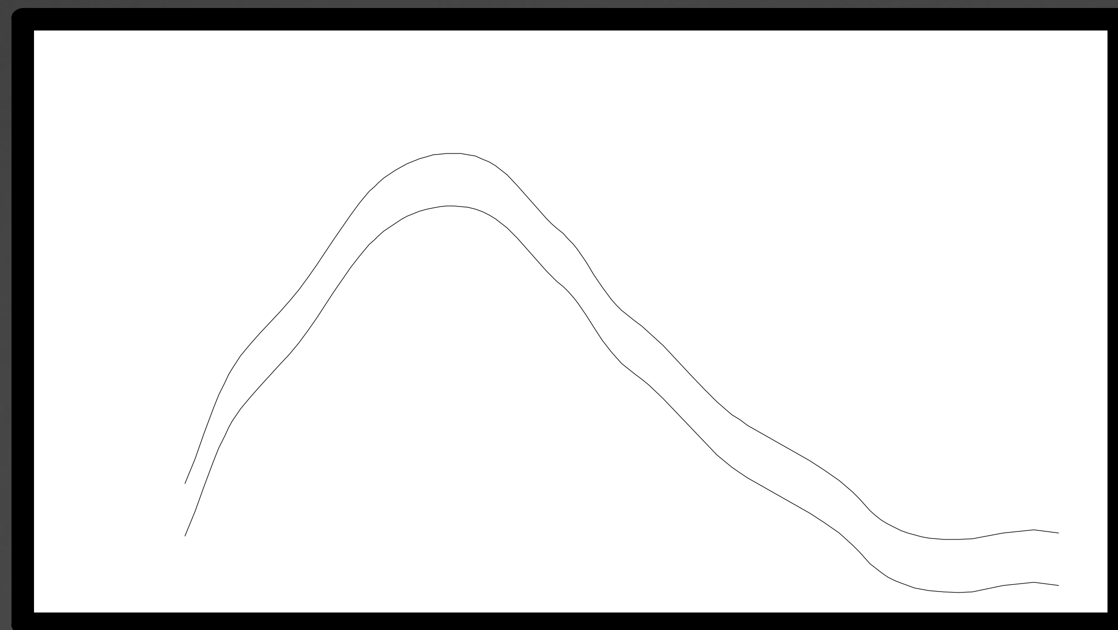
$$\phi_N(k) = \phi(k)^N = (1 - k^2\sigma^2/2)^N$$

$$\phi_N(k) = \underset{N\to\infty}{\to} \exp(-k^2 N\sigma^2/2)$$

Gaussianity also appears under other conditions. Fluctuations of system in equilibrium P ~ exp (- *S*), where the entropy, *S*, is maximum.

Central limit theorem

Intuitive interpretation



Numerical experiment

Covariance and Correlations

Definitions

- $\mathrm{Cov}(X, Y) \stackrel{\mathrm{def}}{=} \mathrm{E}\left[(X - \mu_X)(Y - \mu_Y)\right]$.
- $\mathrm{Cov}(X, Y)$ is denoted by $\sigma_{X,Y}$.
- $\mathrm{Var}(X) = \mathrm{Cov}(X, X)$.
- $\mathrm{Corr}(X, Y) \stackrel{\mathrm{def}}{=} \mathrm{Cov}(X, Y)/\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}$.

Covariance and Correlation

Covariance and Correlation Results (be able to prove any of these).

- $\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$.
- Cauchy-Schwartz Inequality: $[\mathrm{E}(XY)]^2 \le \mathrm{E}(X^2)\mathrm{E}(Y^2)$.
- $\rho_{X,Y} \in [-1,1]$ To proof, use the Cauchy-Schwartz inequality.
- $\mathrm{Cov}(a+bX, c+dY) = bd\,\mathrm{Cov}(X,Y)$.
- $\mathrm{Cov}\left(\sum_i a_i X_i, \sum_i b_i Y_i\right) = \sum_i \sum_j a_i b_j \,\mathrm{Cov}(X_i, Y_j)$. For example, $\mathrm{Cov}(aW + bX, cY + dZ) =$ $ac\,\mathrm{Cov}(W,Y) + ad\,\mathrm{Cov}(W,Z) + bc\,\mathrm{Cov}(X,Y) + bd\,\mathrm{Cov}(X,Z)$.
- $\mathrm{Corr}(a+bX, c+dY) = \mathrm{sign}(bd)\,\mathrm{Corr}(X,Y)$.
- $\mathrm{Var}\left(\sum_i X_i\right) = \sum_i \sum_j \mathrm{Cov}(X_i, X_j) =$ $\sum_i \mathrm{Var}(X_i) + \sum_{i \ne j} \mathrm{Cov}(X_i, X_j)$.
- Parallel axis theorem: $\mathrm{E}(X-c)^2 = \mathrm{Var}(X) + (\mu_X - c)^2$. Hint on proof: first add zero $X - c = (X - \mu_X) + (\mu_X - c)$, then take expectation.

Statistical independence, stationarity & homogeneity

Definition of independence: Continuous random variables *x* and *y* are said to be independent if their joint pdf factors into a product of the marginal pdfs. That is:

$$p(x,y) = p_1(x)p_2(y) \;\; \forall \;\; x \;\&\; y$$

if *x* and *y* are independent then any f(x) and g(y) are also independent.

Show that the convolution of two independent random variables, i.e., $x = x_1 + x_2$, gives: $\phi(k) = \phi_1(k)\phi_2(k)$

Strong Stationarity: if all the statistical properties (i.e. the PDF) of a certain random process are independent of time the process is called strongly stationary.
If only the variance is time independent then the process is weakly stationary.

The Correlation function

Let {x(t)} be a family of random variables indexed by a parameter t (time). We define the following functions:

$$
\begin{aligned}
\mathsf{E}\left(\left[x(t)-\mu(t)\right]\left[x(t')-\mu(t')\right]\right) &= \mathrm{Cov}\left(x(t), x(t')\right) \\
&\equiv \psi(t,t') \\
&\equiv \xi(t,t') - \mu(t)\mu(t')
\end{aligned}
$$

The functions ξ & ψ are called the correlation and covariance function, respectively. s

The correlation function

The formal definition of the correlation function in the limit of continuous random field $f(t)$ is:

$$\xi(t,t') = \langle f(t) f(t') \rangle$$

For a stationary process this simplifies to:

$$\xi(t,t+\tau) = \xi(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(t) f(t+\tau) dt$$

For a multidimensional homogeneous and isotropic random function $f(\vec{x})$ the correlation function satisfies the following condition:

$$\xi(\vec{x}_1, \vec{x}_2) = \xi(|\vec{x}_2 - \vec{x}_1|)$$

The Wiener-Khinchin Theorem

$$
\begin{aligned}
\xi(\tau) &= \int_{-\infty}^{\infty} f(t) f(t+\tau) dt \\
&= \int_{-\infty}^{\infty} \left[\int \tilde{f}(\omega) e^{i\omega t} d\omega \right] \left[\int \tilde{f}^{+}(\omega') e^{-i\omega'(t+\tau)} d\omega \right] dt \\
&= \int_{-\infty}^{\infty} \left| \tilde{f}(\omega) \right|^2 e^{i\omega\tau} d\omega
\end{aligned}
$$

The correlation function is Fourier transform of the power spectrum of the random field.

Principal Component Analysis

Given a vector of data (*e.g.,* galaxy spectrum) with $j=1,...,N_S$ bins which have been measure for $i=1,...,N_G$ objects (galaxies).
PCA is a method that provides a compact description of the data by identifying the linear combination of the input parameters with maximum variance.

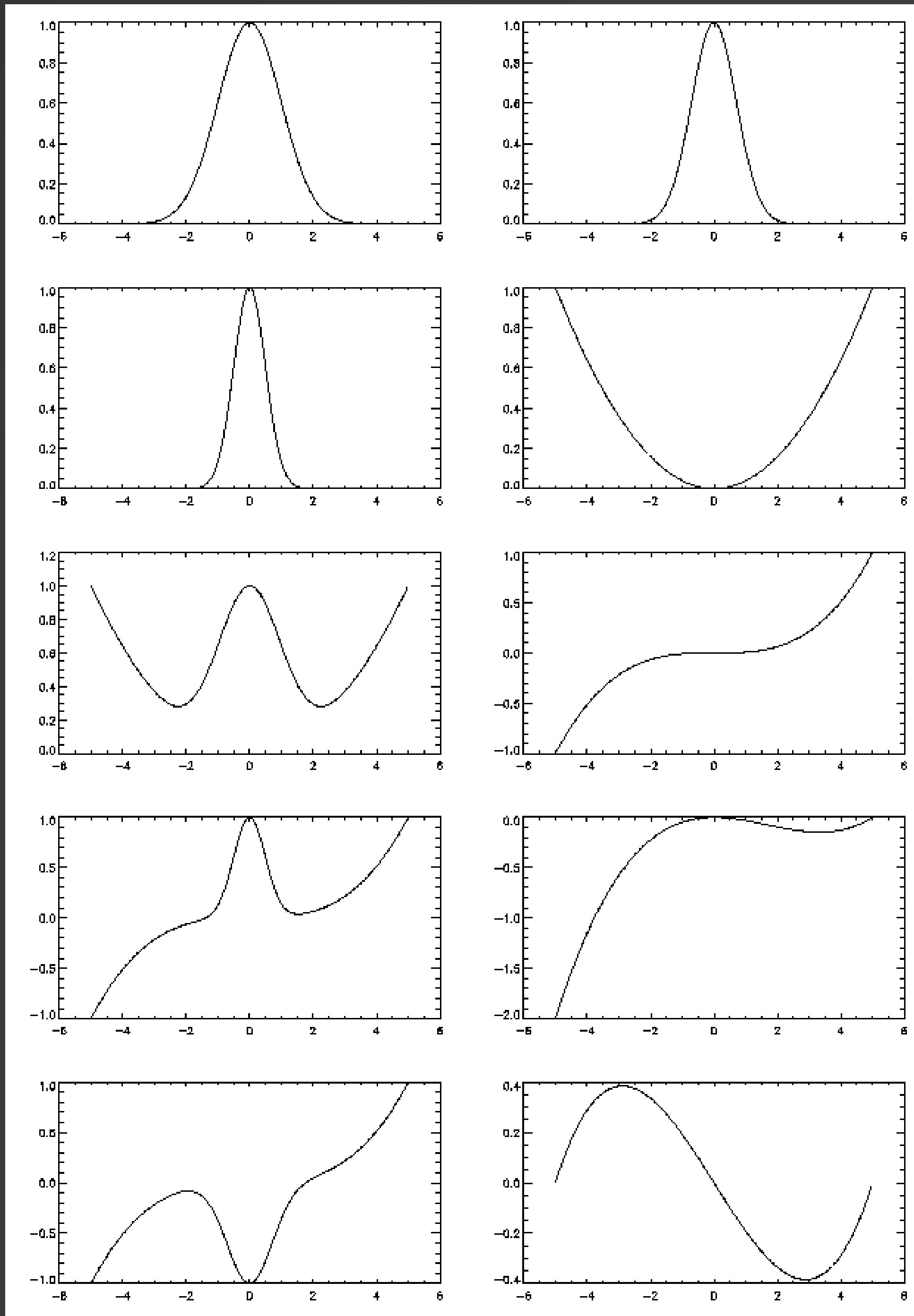
Given the measured values, r_{ij}, where *i* is a given object and *j* runs over the measured attributes of that object. The mean subtracted can be constructed, $X_{ij}=r_{ij}-\langle r_j \rangle$ the covariance matrix of these quantities is:

$$C_{j,k}=\frac{1}{N_G}\sum_{i=1}^{N_G} X_{ij} X_{i,k} \quad 1 \le j \le N_S \quad 1 \le k \le N_S$$

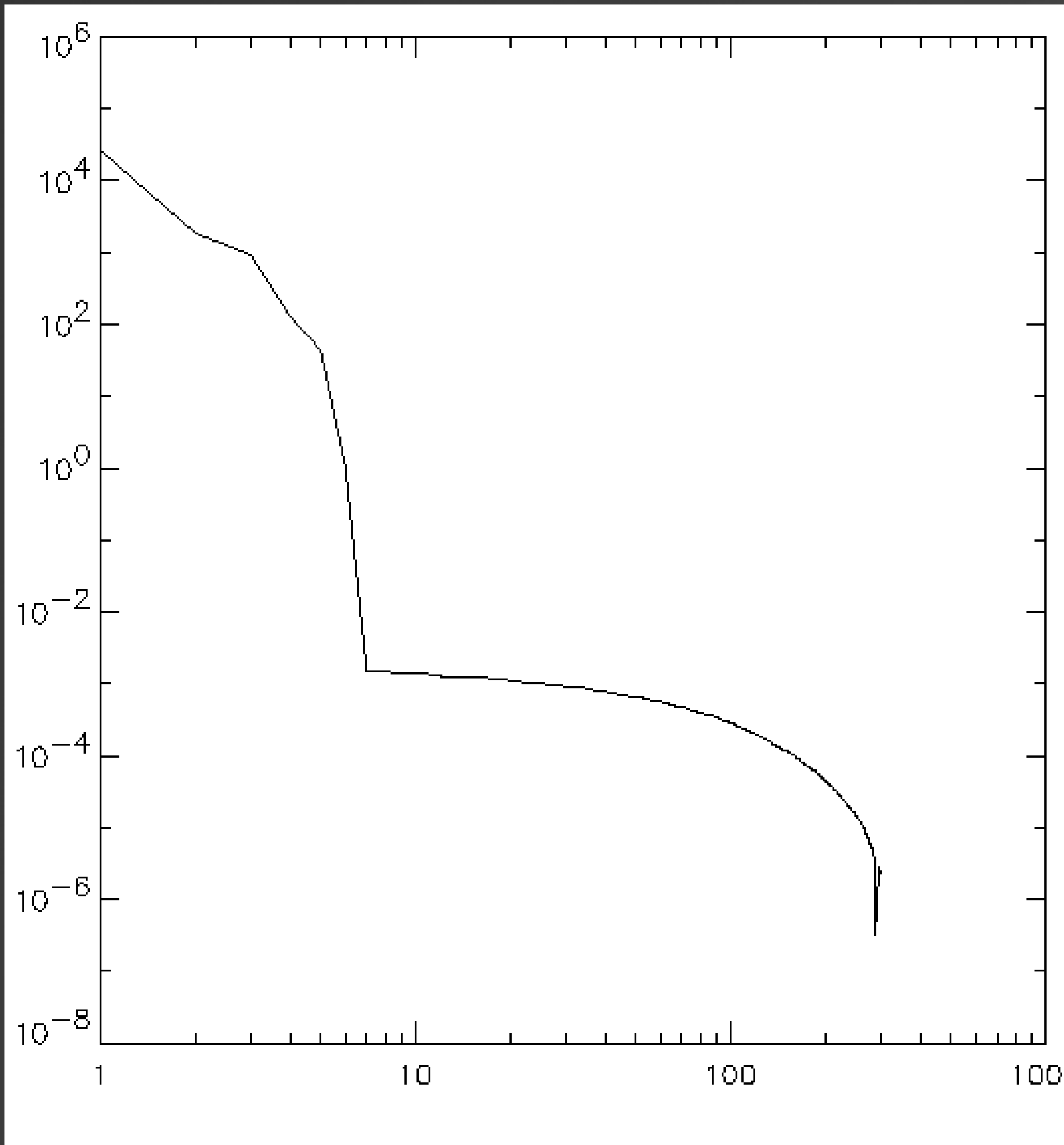
It could be shown that the direction in vector space along which the variance is maximal is the eigenvector $\mathbf{e}_1$ of the matrix equation:

$$\mathbb{C}\,\vec{e}_1=\lambda_1 \vec{e}_1$$

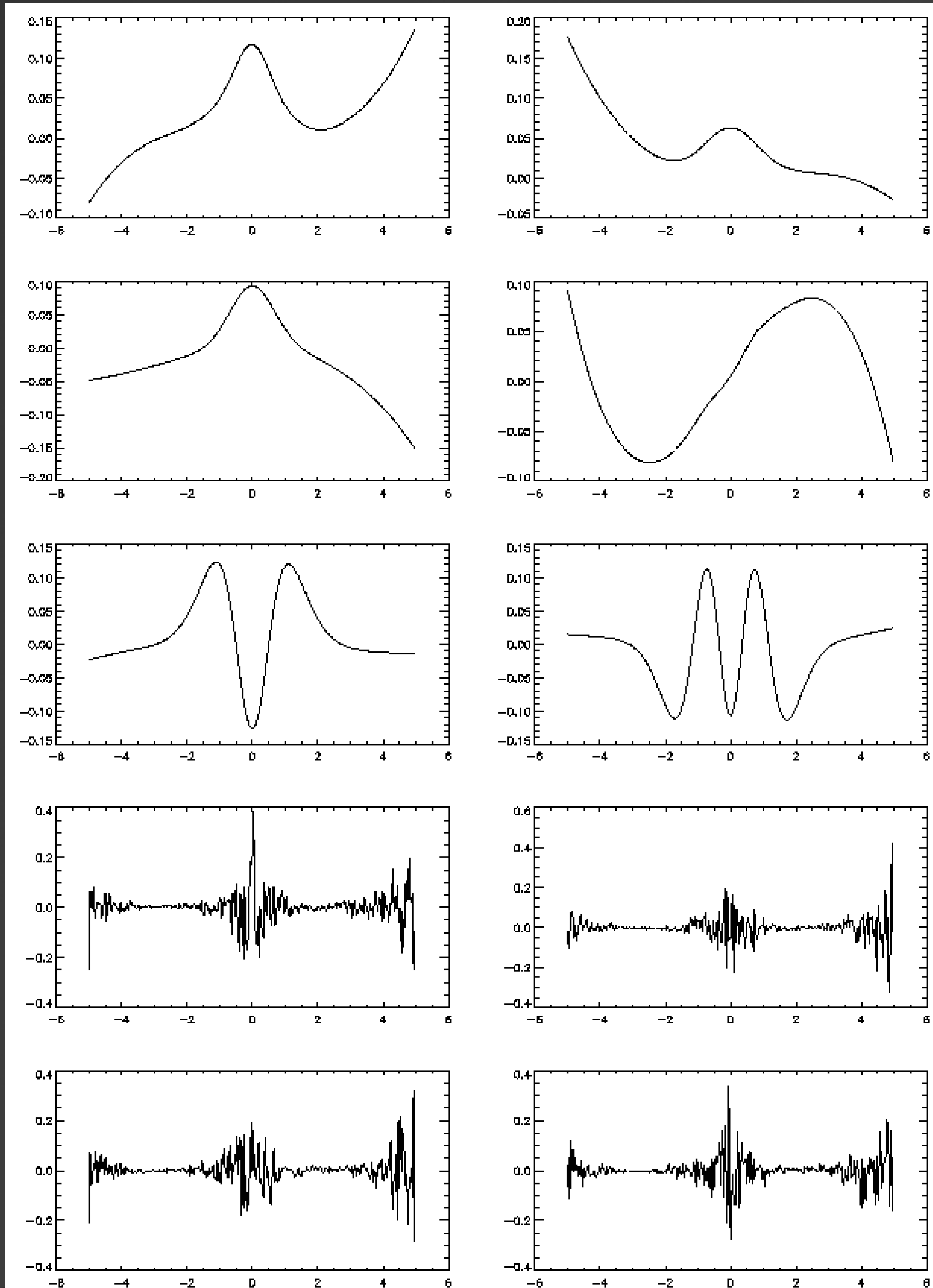
with L_1 is the largest eigenvalue, ...



The input 10 functions that contribute to the 400 measured objects (each has 300 data points)



The eigenvalues of the correlation matrix, the first 7 are much more significant that the others.



The first 10 eigenvectors of the expansion note that they are related to the input functions but no exactly the same, few are rather some are combinations of initial function. Remember that some of the original functions contribute as much as the noise therefore they don't appear as part of the PCAs.

Fast Fourier Transform