Data Science, 2022
Tut 4: Independent Component Analysis

Ex. 1

Exercise: Mixing statistically independent sources
Given some scalar and statistically independent random variables (signals) s_i with zero mean, unit variance,
and a value a_i for the kurtosis that lies between −a and +a, with arbitrary but fixed value of 0 < a. The s_i
shall be mixed like

$$x := \sum_i w_i s_i$$

with constant weights w_i.

- Which constraints do you have to impose on the weights w_i.to guarantee that the mixture has unit variance as well?

Hint

$$\begin{aligned} \mathrm{var}(x) &= \left\langle (x - \langle x \rangle)^2 \right\rangle \\ &= \left\langle x^2 \right\rangle - \langle x \rangle^2 \end{aligned}$$

Tutorial 4

Ex 1.

Given,

* Scalar & statistically independent random variables
* mean $= 0$
  variance $= 1$.

a value for kurtosis: $-a$ to $+a$
where, $0<a$

s_i is mixed like.

$$x := \sum_i w_i s_i \qquad -(1)$$

$w_i \rightarrow$ constraint weight

Q. which constraints do you have to impose on the weights w_i, to guarantee that the mixture has unit variance as well?

Hint:

$$var(x) = \langle\langle x - \langle x\rangle\rangle^2\rangle$$
$$= \langle x^2\rangle - \langle x\rangle^2$$

Continuing on the hint, substituting (1)

$$var(x) = \langle(\sum_i w_i s_i)^2\rangle - \langle \sum_i w_i s_i\rangle^2$$

$$= \langle(\sum_i w_i s_i)^2\rangle - \left(\sum_i w_i \langle s_i\rangle\right)^2$$

$$= \left\langle \left(\sum_i w_i s_i\right)\left(\sum_j w_j s_j\right)\right\rangle - \left(\sum_i w_i \langle s_i \rangle\right)\left(\sum_j w_j \langle s_j \rangle\right)$$

$$= \left\langle \sum_{i,j} w_i w_j s_i s_j \right\rangle - \sum_{i,j} w_i w_j \langle s_i \rangle \langle s_j \rangle$$

$$= \sum_{i,j} w_i w_j \langle s_i s_j \rangle - \sum_{i,j} w_i w_j \langle s_i \rangle \langle s_j \rangle$$

$$= \sum_i w_i w_i \left(\langle s_i s_i \rangle - \langle s_i \rangle \langle s_i \rangle\right) + \sum_{i,j : i \neq j} w_i w_j \left(\langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle\right)$$

$$= \sum_i w_i^2 \underbrace{\left(\langle s_i s_i \rangle - \langle s_i \rangle^2\right)}_{= var(s_i) = 1} + \sum_{i,j : i \neq j} w_i w_j \underbrace{\left(\langle s_i \rangle \langle s_j \rangle - \langle s_i \rangle \langle s_j \rangle\right)}_{= 0}$$

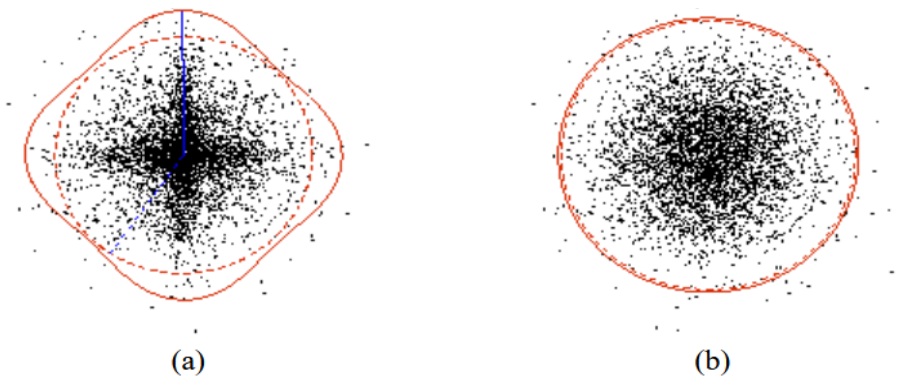
$$= \sum_i w_i^2$$

$\therefore$ the constraint is

$$\underline{\underline{\sum_i w_i^2 = 1}}$$

Ex.2

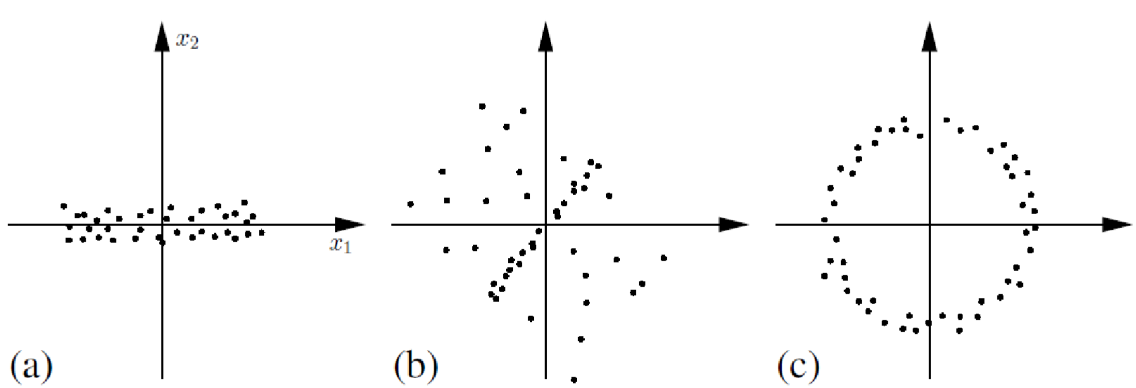
Two examples of joint probability densities are shown in following figure. One is a mixture of arbitrary non-Gaussian densities, and the other one a mixture of Gaussians. The dashed curves around the densities plot the projected variance measured in all directions. The dashed line marks the direction of maximum variance, that is, the first principal component. Similarly, the values of kurtosis are shown using solid curves and the direction of maximum kurtosis with a solid line.



Example joint propability densities. (a) For non-Gaussian densities the principal (dashed line) and independent (solid line) directions can be identified, where as (b) for Gaussian ones the directions are all equal. The corresponding dashed and solid curves show the values of variance and kurtosis in all directions respectively.

Referring to above provide the guess independent components and distributions from data

- Decide whether the following distributions can be linearly separated into independent components. If yes,
- sketch the (not necessarily orthogonal) axes onto which the data must be projected to extract the independent components. Draw these axes also the marginal distributions of the corresponding components.



Ex2.

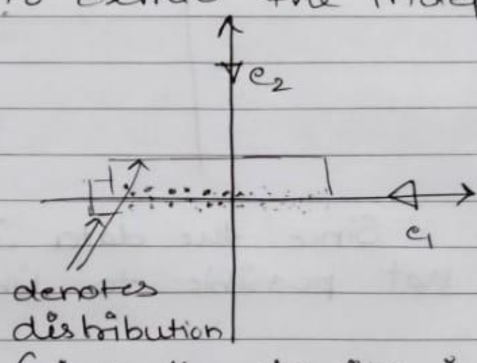
* Need to guess independent components & distribution.

1) Decide whether the following distributions can be linearly separated into independent components. If yes,

2) Sketch the axes onto which the data must be projected to extract the independent components. Draw these axes also the marginal distributions of the corresponding components

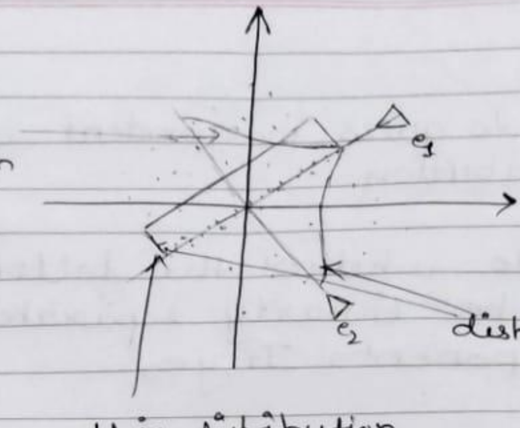
let us have e_1 & e_2 as vectors that help to extract the independent components

a)



(since the density is almost even in both directions) b)

this distribution is denser towards the center



distribution

this distribution is pretty much even across.

c)

Since the data is circular it is not possible to linearly separate it.