

INDEPENDENT COMPONENT ANALYSIS

Independent Component Analysis is a statistical generative model which aims to reveal hidden independent additive components from multi-dimensional data.

1 Mathematical Preliminaries

1.1 Single Random Variable

Let the Probability Mass Function of a discrete random variable X be defined as follows.

	$x = 1$	$x = 2$	$x = 3$	$x = 4$
$P(X = x)$	$1/6$	$1/3$	$1/3$	$1/6$

The expectation $E(X)$ of this random variable is defined as follows.

$$\begin{aligned} E(X) &= \sum_k x_k P(X = x_k) \\ &= 1 \times \frac{1}{6} + 2 \times \frac{1}{3} + 3 \times \frac{1}{3} + 4 \times \frac{1}{6} \\ &= \frac{15}{6} \end{aligned}$$

If a new variable $Y = 2X + 1$ is defined, the PMF of Y becomes

	$y = 2 \times 1 + 1$	$y = 2 \times 2 + 1$	$y = 2 \times 3 + 1$	$y = 2 \times 4 + 1$
$P(Y = y = 2x + 1)$	$1/6$	$1/3$	$1/3$	$1/6$

$$\begin{aligned}
E(Y) &= \sum_k y_k P(Y = y_k) \\
&= \sum_k (2x_k + 1) P(Y = y_k) \\
&= \sum_k (2x_k + 1) P(X = x_k) \\
&= 2E(X) + 1
\end{aligned}$$

Hence, expectation is a linear operator which satisfies $E(aX + b) = aE(X) + b$.

The variance of a discrete random variable X is defined as

$$\begin{aligned}
Var(X) &= E((X - E(X))(X - E(X))) \\
&= E(X^2 - 2E(X)X + E(X)E(X)) \\
&= E(X^2) - 2E(X)E(X) + E(X)E(X) \\
&= E(X^2) - E(X)^2
\end{aligned}$$

In case of continuous random variables, instead of Probability Mass Function a Probability Density Function is defined because the sample space is continuous and infinite as opposed to discrete and finite.

The expectation and variance of a continuous random variable are defined as

$$\begin{aligned}
E(X) &= \int xP(X = x)dx \\
Var(X) &= \int (x - E(X))^2 P(X = x)dx
\end{aligned}$$

Just as in the case of discrete random variables where the sum of probabilities of all events in sample space must sum to 1, the area under PDF curve in the case of continuous random variable must also integrate to 1.

An example of a well known continuous random variable is the Gaussian random variable whose PDF is defined as follows.

$$P(X = x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

with the following expectation and variance.

$$\begin{aligned}
E(X) &= \mu \\
Var(X) &= \sigma^2
\end{aligned}$$

1.2 A Pair of Random Variables

Let the joint Probability Mass Function of a pair of discrete random variables X and Y be defined as follows.

$P(X=x, Y=y)$	$x = 1$	$x = 2$	$x = 3$
$y = 1$	$3/8$	$3/16$	$3/16$
$y = 2$	$1/8$	$1/16$	$1/16$

As usual, the sum of all probabilities is equal to 1.

Given the joint distribution of a set of random variables, one can also find out the distribution of a subset of those random variables and the distribution of such a subset is called the marginal distribution. In the above example, the marginal distribution of X and Y can be computed by summing up each column and row respectively in the table above.

$$P(X = x_k) = \sum_{y_l} P(X = x_k, y = y_l)$$

$$P(Y = y_k) = \sum_{x_l} P(X = x_l, y = y_k)$$

In the case of a pair of continuous random variables, their joint Probability Distribution Function becomes a surface in three dimensions. The marginal distributions take the following formulae.

$$P(X = x) = \int_y P(X = x, y = y) dy$$

$$P(Y = y) = \int_x P(X = x, y = y) dx$$

Two random variables are said to statistically independent if and only if

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

For independent random variables, the following result holds.

$$\begin{aligned}
E(f(X)g(Y)) &= \int \int f(x)g(y)P(X = x, Y = y)dx dy \\
&= \int \int f(x)g(y)P(X = x)P(Y = y)dx dy \\
&= \int f(x)P(X = x)dx \int g(y)P(Y = y)dy \\
&= E(f(X))E(g(Y))
\end{aligned}$$

A weaker form of independence is uncorrelatedness. If $f(X)$ and $g(Y)$ are defined such that $f(X) = X$ and $g(Y) = Y$ and $E(XY) = E(X)E(Y)$ holds true, then two variables are said to be uncorrelated.

1.3 Sum of Random Variables

If X denotes the outcome of roll of a dice and Y denotes the outcome of roll of another dice, then $Z = X + Y$ denotes a random variable whose value is equal to the sum of the two rolls. If PMFs of X and Y are known, what is the PMF of $Z = X + Y$?

$$\begin{aligned}
P(Z = X + Y = z) &= \sum_x P(X = x, Y = z - x) \\
&= \sum_y P(X = z - y, Y = y)
\end{aligned}$$

It is easy to see why the above equation holds. To find $P(Z = z)$, one has to sum the probabilities $P(X = x, Y = y)$ of all possible scenarios where $x + y = z$. In the case of continuous random variables the formulae become

$$\begin{aligned}
P(Z = X + Y = z) &= \int P(X = x, Y = z - x)dx \\
&= \int P(X = z - y, Y = y)dy
\end{aligned}$$

The expectation of a sum of two random variables is the sum of expectations of individual random variables.

$$\begin{aligned}
E(X + Y) &= \int \int (x + y)P(X = x, Y = y)dx dy \\
&= \int \int xP(X = x, Y = y)dx dy + \int \int yP(X = x, Y = y)dx dy \\
&= \int x \int P(X = x, Y = y)dy dx + \int y \int P(X = x, Y = y)dx dy \\
&= \int xP(X = x)dx + \int yP(Y = y)dy \\
&= E(X) + E(Y)
\end{aligned}$$

Interestingly, If X and Y are independent Gaussian random variables then the above result holds for variances too.

$$\begin{aligned}
E(X + Y) &= E(X) + E(Y) \\
Var(X + Y) &= Var(X) + Var(Y)
\end{aligned}$$

2 ICA Motivation