Formulaire

NICOLAS LE HIR

TABLE DES MATIÈRES

1 DÉCOMPOSITION BINAIRE

Result: Entier n en binaire
$L \leftarrow$ liste vide [];
$r \leftarrow 0$;
while $n > 0$ **do**
  $r \leftarrow n\%2$;
  $l \leftarrow l + [r]$;
  $n \leftarrow (n - r)/2$;
end
$L \leftarrow \text{reversed}(L)$;
return L

Algorithm 1: Binary decomposition of integer n

2 DISTANCES

Here are some common distances in $\mathbb{R}^2$ and $\mathbb{R}^3$.

2.1 Distances in two dimensions

We consider two points M_1 and M_2 in the 2D space $\mathbb{R}^2$ with coordinates (x_1, y_1), and (x_2, y_2), respectively.

L2

$$d(M_1, M_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{1}$$

L1

$$d(M_1, M_2) = |x_1 - x_2| + |y_1 - y_2| \tag{2}$$

L∞

$$d(M_1, M_2) = \max(|x_1 - x_2|, |y_1 - y_2|) \tag{3}$$

weighted L1 :
let α_1 and α_2 be real numbers ($\in \mathbb{R}$).

$$d(M_1, M_2) = \alpha_1|x_1 - x_2| + \alpha_2|y_1 - y_2| \tag{4}$$

2.2 Distances in three dimensions

We consider two points M_1 and M_2 in the 3D space $\mathbb{R}^3$ with coordinates (x_1, y_1, z_1), and (x_2, y_2, z_2), respectively.

L2

$$d(M_1, M_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \tag{5}$$

L1

$$d(M_1, M_2) = |x_1 - x_2| + |y_1 - y_2| + |z_1 - z_2| \tag{6}$$

L∞

$$d(M_1, M_2) = \max(|x_1 - x_2|, |y_1 - y_2|, |z_1 - z_2|) \tag{7}$$

weighted L1 :
let α_1, α_2 and α_3 be real numbers ($\in \mathbb{R}$).

$$d(M_1, M_2) = \alpha_1|x_1 - x_2| + \alpha_2|y_1 - y_2| + \alpha_3|z_1 - z_2| \tag{8}$$

3 LIKELIHOOD / VRAISEMBLANCE

— Observations : $(x_1, \dots, x_n)$

— Model : p (for instance a normal law)

— Parameters : θ (for instance (μ, σ), the mean and the standard deviation of the normal law)

The likelihood writes :

$$L(\theta) = p(x_1, \dots, x_n|\theta) \tag{9}$$

4 DERIVATIVE / DÉRIVÉE

Let $f : \mathbb{R} \to f(\mathbb{R})$ be a real function.
When it exists, the derivative of f in x is defined by :

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{10}$$

Examples :

If $g : x \mapsto 3x$, then the derivative exists and $\forall x \in \mathbb{R}, x, g'(x) = 3$

If $h : x \mapsto x^2$, then the derivative exists and $\forall x \in \mathbb{R}, h'(x) = 2x$.

If $h : x \mapsto |x|$, then the derivative exists only if $x \neq 0$.

5 EXPECTED VALUE / ÉSPÉRANCE

Let X be a discrete random variable that takes the values x_i with probability p_i.

The **expected value** of X writes

$$E(X) = \sum_{i=1}^{n} p_i x_i \tag{11}$$

Example :
If X is a constant random variable : $X = \alpha$

$$\sum_{i=1}^{n} p_i x_i = \sum_{i=1}^{n} p_i \alpha = \alpha \sum_{i=1}^{n} p_i \tag{12}$$

6 K-MEANS / K MOYENNES

- Datapoints $(x_1, \ldots, x_n)$
- Centroids $(c_1, \ldots, c_n)$ (one centroid per point, however the number of different centroids is smaller than the number of datapoints)

The inertia I is given by :

$$I = \sum_{i=1}^{n} d(x_i, c_i)^2 \tag{13}$$

7 ENTROPY / ENTROPIE

The entropy of a discrete random variable X that takes the values x_i with probability p_i is given by :

$$H(X) = -\sum_{i=1}^{n} p_i \log(p_i) \tag{14}$$

Examples :

- Entropy of certain distribution.

$$H = 0 \tag{15}$$

- Entropy of uniform distribution with n values :

$$\begin{aligned} H &= -\sum_{i=1}^{n} \frac{1}{n} \log \frac{1}{n} \\ &= -n \times \frac{1}{n} \times \log \frac{1}{n} \\ &= \log n \end{aligned} \tag{16}$$

8 COMPLEXITY

Let n be the size of the problem.
Polynomial complexity :

$$a_k n^k + a_{k-1} n^{k-1} + \cdots + n \tag{17}$$

Exponential complexity :

$$k^n \tag{18}$$

with $k > 1$