Any Single Gaussian

Tuesday, August 24, 2021 6:08 PM

Julen

How to summarize data?

$x^1, x^2, \ldots \ldots x^m.$

- mean $\mu = \frac{1}{m} \sum_i x^i$
- Variance $\sigma^2 = \frac{1}{m} \sum_i (x^i - \mu)^2$

What about in higher dim? mean is still average.

Variance?

Variance along direction (unit vector v):

Variance of $x^1 \cdot v,\ x^2 \cdot v, \ldots \ldots x^m \cdot v$

$$\sigma_v^2 = \frac{1}{m} \sum_i (x^i \cdot v - \mu \cdot v)^2$$

$$= \frac{1}{m} \sum_i ((x^i - \mu) \cdot v)^2$$

$$= v^T \underbrace{\left(\frac{1}{m} \sum_i (x^i - \mu)(x^i - \mu)^T\right)}\, v$$

Σ : covariance matrix

$$\Sigma_{ii} = \mathbb{E}\left((X_i - \mu_i)^2\right)$$

$$\Sigma_{ij} = \mathbb{E}\left((X_i - \mu_i)(X_j - \mu_j)\right).$$

Probability distribution $\quad Pr(X = a)$

Probability density function $\quad pdf(X = a)$
(continuous densities)

e.g. uniform in $[a, b]$

$$p(x = t) = \begin{cases} \frac{1}{|b-a|} & t \in [a,b] \\ 0 & t \notin [a,b] \end{cases}$$

Can have $p(x) = f(x)$ where

$f(x) \geq 0$

nonnegative, integrable. $\quad \int f < \infty.$

Most important example

$\boxed{\text{Gaussian Density.}}$ $\quad$ 1-dim $\quad e^{-\frac{x^2}{2}}$

Gaussian Density.

1-dim

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Q. Why $\frac{1}{\sqrt{2\pi}}$?

— $x \in \mathbb{R}^d$ $\quad p(x) = \frac{1}{(\sqrt{2\pi})^d} e^{-\frac{\|x\|^2}{2}}$

$\|x\|^2 = \sum_i x_i^2$ $\quad = \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}}$

Product of 1-dim Gaussians along each coordinate.

F1. for any unit vector v, $x \sim N(0, I_d)$

$x \cdot v \sim N(0, 1)$

$$\underset{=t}{p(x \cdot v)} = \int_{x:\, x \cdot v = t} \frac{1}{(\sqrt{2\pi})^d} e^{-\frac{\|x\|^2}{2}} dx$$

Write x in the basis $\{v, v_2, \ldots v_d\}$

$\|x\|^2 = \sum (x \cdot v_i)^2$

$$e^{-\frac{t^2}{2}} \int \frac{1}{\cdots} e^{-\frac{\sum_{i=2}^{d} (x \cdot v_i)^2}{2}}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \int \frac{1}{(\sqrt{2\pi})^{d-1}} e^{-\frac{\sum_{i=2} \langle x, v_i \rangle^2}{2}}$$

$x \cdot v_1 = t$

$y = (x\cdot v_2 \dots x \cdot v_d)$

$dy = \left|\det[v_2 \dots v_d]\right| dx$

$$= \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} \int_y \frac{1}{(\sqrt{2\pi})^{d-1}} \cdot e^{-\frac{\|y\|^2}{2}} dy$$

$$= \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}.$$

F2. $\frac{1}{\sqrt{2\pi}} \int e^{-\frac{x^2}{2}} dx = 1$

Let $p(x)$ " $p(x,y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$

$$\frac{1}{2\pi} \int e^{-\frac{(x^2+y^2)}{2}} dx\,dy = \frac{1}{\cancel{2\pi}} \int_{r=0}^{\infty} \cancel{2\pi} r \cdot e^{-\frac{r^2}{2}} dr$$

$u = \frac{r^2}{2}$ $du = r\,dr$

$$= \int_u e^{-u} du = 1.$$

$$\int p(x,y) = \left(\int p(x)\right)^2 \Rightarrow \int p(x) = 1.$$

General Gaussian $N(\mu, \sigma^2)$

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

Multivariate $$N(\mu, \Sigma) = \frac{1}{(\sqrt{2\pi})^d |\det(\Sigma)|} e^{-\frac{(x-\mu)^T \Sigma^{-1} (x-\mu)}{2}}$$

$$X \longrightarrow Ax+b$$

$$N(0, I_d) \longrightarrow N(b, AA^T)$$

$$X \sim N(\mu, \Sigma) \Rightarrow \Sigma^{-\frac{1}{2}}(x-\mu) \sim N(0, I_d)$$

P1. Estimate Gaussian given random iid samples.

$x^1, x^2, \ldots\ldots x^m.$

Natural choices: $\tilde{\mu} = \frac{1}{m} \sum_i x^i$

$$\tilde{\Sigma} = \frac{1}{} \sum (x^i - \tilde{\mu})(x^i - \tilde{\mu})^T$$

$$\tilde{\Sigma} = \frac{1}{m} \sum_i (x^i - \tilde{\mu})(x^i - \tilde{\mu})'$$

Something better?

<u>Goal.</u> Find $\tilde{\mu}$, $\tilde{\Sigma}$ s.t.

likelihood $N(\tilde{\mu}, \tilde{\Sigma})$ generates $x^1 \dots x^m$

is maximized.

<u>Thm.</u> Empirical mean, covariance give max likelihood estimates.

<u>Pf.</u> Consider the case when $\Sigma = \sigma^2 I$.

$$P(\text{data}) = \prod_{j=1}^{m} \frac{1}{(\sqrt{2\pi}\,\sigma)^d} e^{-\frac{\|x^j - \mu\|^2}{2\sigma^2}}$$

$$\log(P(\text{data})) = -\log(\sqrt{2\pi}\,\sigma) - \sum_j \frac{\|x^j - \mu\|^2}{2\sigma^2}$$

$$\frac{\partial}{\partial \sigma} = -\frac{md}{\sigma} + \frac{1}{\sigma^3} \sum_j \|x^j - \mu\|^2 = 0$$

$$\sum \|x^j - \mu\|^2$$

$$\Rightarrow \quad \sigma^2 = \frac{1}{d} \cdot \frac{1}{m} \sum_j \| x^j - \mu \|^2$$

$$\frac{\partial L}{\partial \mu_i} = 0 \quad \Rightarrow$$

$$-\frac{1}{2\sigma^2} \cdot 2 \sum_j (x_i^j - \mu_i) = 0$$

$$m \mu_i = \sum x_i^j$$

$$\mu_i = \frac{1}{m} \sum x_i^j$$

$$\mu = \frac{1}{m} \sum x^j .$$

<u>Thm</u>. [Central Limit Theorem] $X, X, \ldots X$ iid random variables from some distribution with bounded variance $\mathrm{Var}(X_1) < \infty$.

Let $Y_n = \frac{1}{n} \sum_i X_i$

Then $Y_n \longrightarrow N\left(\mathbb{E}(X_1), \frac{1}{n} \mathrm{Var}(X_1)\right)$

Then $Y_n \longrightarrow N\left(E(X_1), \frac{1}{n} \mathrm{Var}(X_1)\right)$

(converges in distribution)

This can be made more qualitative.

<u>Thm.</u> [Berry-Esseen] $X_1 \dots X_n$ independent R.V.

$Y_n = \sum_{i=1}^{n} X_i$. $Z_n \sim N(\mathbb{E}(Y_n), \mathrm{Var}(Y_n))$.

Then, for any $t \in \mathbb{R}$,

$$\left| \Pr(Y_n \le t) - \Pr(Z_n \le t) \right| \le c \cdot \frac{\sum_{i=1}^{n} E(|X_i|^3)}{\mathrm{Var}(Y_n)^{3/2}}$$

$c \in [0, 1]$.

<u>Example</u> $X_i = \begin{cases} -1 & \text{w.p. } \frac{1}{2} \\ 1 & \text{w.p. } \frac{1}{2} \end{cases}$

$E(|X_i|^3) = 1$

$\mathrm{Var}(Y_n) = n \mathrm{Var}(X_1) = n \cdot \left(\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 1\right) = n$

$\therefore \forall t \qquad Z_n \sim (0, n)$

$$\left| \Pr(Y_n \le t) - \Pr(Z_n \le t) \right| \le c \cdot \frac{n}{n^{3/2}} \le \frac{c}{\sqrt{n}}.$$

"Invariance" Principle.

Note, bound also holds for $\Pr(Y_n > t)$

$$\left| \Pr(Y_n > t) - \Pr(Z_n > t) \right| = \left| (1 - \Pr(Y_n \le t)) - (1 - \Pr(Z_n \le t)) \right|$$
$$= \left| \Pr(Y_n \le t) - \Pr(Z_n \le t) \right|.$$
