3 Probability

3.1 Basic Properties

For events E_1 and E_2, if they are disjoint, i.e. $E_1 \cap E_2 = \emptyset$, then $\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2)$

Definition 4. (Conditional probability) For events A and B, and $\mathbb{P}(A) > 0$,

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$$

We can define the conditional expectation as

$$\mathbb{E}\left[Y|X=x\right] \triangleq \sum_{y \in \mathcal{Y}} y \cdot p\left(Y=y|X=x\right)$$

Definition 5. (Covariance) For two random variables X and Y, the covariance is defined by

$$\mathrm{Cov}\left[X,Y\right] = \mathbb{E}\left[XY\right] - \mathbb{E}\left[X\right]\mathbb{E}\left[Y\right]$$

When the covariance of X and Y is 0, we call them **uncorrelated variables**.

Definition 6. (Independent) For two random variables, when the joint pdf can be written as the product of two RVs' pdf

$$f\left(x,y\right) = f_X\left(x\right) f_Y\left(y\right),$$

we call them **independent**.

Theorem 2. We have:

∘ (Multiplication Rule) For events A and B,

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B|A) = \mathbb{P}(B)\mathbb{P}(A|B);$$

∘ (Total probability rule) $B_1, B_2, \ldots, B_k$ form a partition of Ω, $\forall i \neq j, B_i \cap B_j = \emptyset, \cup_{i=1}^{k} B_i = \Omega$, we have:

$$\mathbb{P}(A) = \sum_{i=1}^{k} \mathbb{P}(B_i)\mathbb{P}(A|B_i);$$

∘ (Bayes Rule)

$$\mathbb{P}(B_1|A) = \frac{\mathbb{P}(A \cap B_1)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_1)\mathbb{P}(B_1)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_1)\mathbb{P}(B_1)}{\sum_{i=1}^{k} \mathbb{P}(A|B_i)\mathbb{P}(B_i)}.$$

3.2 Gaussian Distribution

3.2.1 Normal Distribution

• If random variable $X \in \mathbb{R}$, $X \sim \mathcal{N}(\mu, \sigma^2)$, where $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}$, then the density function of it is:

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

• $\mathbb{E}[X] = \mu$; $\mathrm{var}(X) = \sigma^2$.

3.2.2 Multivariate Gaussian Distribution

• If random variable $\boldsymbol{X} \in \mathbb{R}^n$, $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} \in \mathbb{R}^n$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$ is symmetric and positive semi-definite (PSD), then the density function of it is:

$$p(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

• $\mathbb{E}[\boldsymbol{X}] = \boldsymbol{\mu}$; $\mathrm{cov}(\boldsymbol{X}) = \boldsymbol{\Sigma}$.

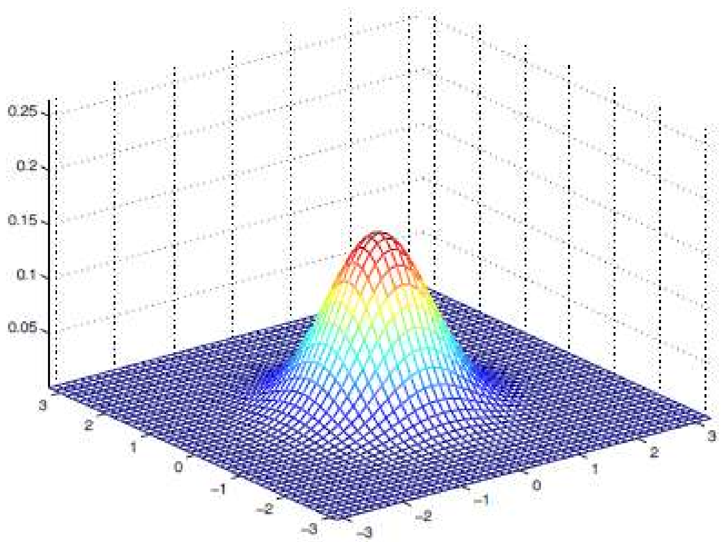

Figure 1: Multivariate Gaussian's p.d.f