Gaussian Processes

$X(t)$ Gaussian Process: $\forall n, \; \forall t_1 \cdots t_n$ $\quad X = (X(t_1), \cdots, X(t_n))^T, \quad X \sim N(\mu, \Sigma).$ $\begin{cases} \mu = E[X] \\ \Sigma = E[(X-\mu)(X-\mu)^T] \end{cases}$

$n = 1, \quad X \sim N(\mu, \sigma^2). \quad f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

$n = 2, \quad X \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho), \quad f_{X_1 X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right)\right]$

Matrix-Vector Notation. $X \in \mathbb{R}^n, \quad X \sim N(\mu, \Sigma). \quad \mu \in \mathbb{R}^n, \quad \Sigma \in \mathbb{R}^{n \times n}.$

$$f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det \Sigma)^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right]$$

Examine $f_X(x)$ is indeed a distribution: $\quad f_X(x) \geq 0, \quad \int_{\mathbb{R}^n} f_X(x)\,dx = 1$

i) $f_X(x) \geq 0.$ $\quad \Sigma$ pos. def. $\Rightarrow \det\Sigma > 0, \quad f_X(x) \geq 0$ $\quad \Rightarrow ? \; \frac{1}{2}\sum \alpha_k x_k^2 + \sum x_i x_j k_{ij}$

ii) $\int_{\mathbb{R}^n} f_X(x)\,dx = 1, \quad \int_{\mathbb{R}^n} f_X(x)\,dx = \frac{1}{(2\pi)^{\frac{n}{2}}(\det\Sigma)^{\frac{1}{2}}} \int_{\mathbb{R}^n} \exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right]dx$

Diagonalize Σ: Σ Covariance Matrix, $\Sigma = \Sigma^T$. Σ pos. def.

$\begin{cases} \Sigma = U^T \Lambda U \; (\text{Eigen-Decomposition}) \\ UU^T = U^T U = I \end{cases}$

$\therefore \Sigma^{-1} = (U^T \Lambda U)^{-1} = U^{-1} \Lambda^{-1} (U^T)^{-1} = U^T \Lambda^{-1} U$

$\lambda_1 \sim \lambda_n \geq 0 \Rightarrow \sqrt{\lambda_k}$ exists

$= U^T \Lambda^{-\frac{1}{2}} \Lambda^{-\frac{1}{2}} U = B^T B$

$\therefore \frac{1}{(2\pi)^{\frac{n}{2}}(\det\Sigma)^{\frac{1}{2}}} \int_{\mathbb{R}^n} \exp\left[-\frac{1}{2}\underbrace{(x-\mu)^T B^T}_{Y^T}\underbrace{B(x-\mu)}_{Y}\right] dx \quad \xmapsto{Y = B(x-\mu)} \quad \frac{1}{(2\pi)^{\frac{n}{2}}(\det\Sigma)^{\frac{1}{2}}} \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}Y^T Y\right) \underbrace{(\det\Sigma)^{\frac{1}{2}}}_{①\;②} dY$

② Jacobian

$\begin{cases} dx = \frac{\partial x}{\partial Y} dY \\ \frac{\partial Y}{\partial x} = B \end{cases}$ $\quad dx = |\det B|^{-1} dY = \left[\left(\prod_k \lambda_k\right)^{-\frac{1}{2}}\right]^{-1} dY = \left(\prod_k \lambda_k\right)^{\frac{1}{2}} dY = (\det\Sigma)^{\frac{1}{2}} dY$

$\underbrace{= \det(\Sigma^{-1}) = \det(U) \cdot \det(\Lambda^{-1}) \det(U^T) = \det(\Lambda^{-1}) = \left(\prod_k \lambda_k\right)^{-1}}_{[\det(B)]^2}$

③ Integration Range: still $\mathbb{R}^n$

$\therefore \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\sum_k y_k^2\right) dy_1 \cdots dy_n = \prod_k \underbrace{\left[\frac{1}{(2\pi)^{\frac{1}{2}}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}y_k^2\right) dy_k\right]}_{=1}$

$= 1 \quad \longrightarrow \text{is a distribution}$

$I = \int_{\mathbb{R}} \exp\left(-\frac{y^2}{2}\right) dy$

$I^2 = \int_{\mathbb{R}^2} \exp\left(-\frac{x^2+y^2}{2}\right) dx\,dy \quad$ In polar coord.

$I^2 = \int_0^{2\pi}\int_0^\infty \exp\left(-\frac{\rho^2}{2}\right)\rho\,d\rho\,d\theta = 2\pi \cdot \int_0^\infty \exp\left(-\frac{\rho^2}{2}\right)\rho\,d\rho$

$= 2\pi$

$\Rightarrow I = (2\pi)^{\frac{1}{2}}$

n-dimensional Characteristic Function:

$$X \in \mathbb{R}^n \text{ r.v.} \quad \phi_X(w) = E[\exp(jw^T X)]$$

$$w = (w_1, \cdots, w_n)^T$$

$$X \sim N(\mu, \Sigma), \quad \phi_X(w) = \int_{\mathbb{R}^n} \exp(jw^T z) \cdot f_X(x)\,dx = \frac{1}{(2\pi)^{n/2}(\det\Sigma)^{1/2}} \int_{\mathbb{R}^n} \exp(jw^T x)\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)dx$$

Informal : 1-Dim

$$-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu) + jw^T x \xrightarrow{\text{1-Dim}} -\frac{1}{2\sigma^2}(x-\mu)^2 + jwx = -\frac{1}{2\sigma^2}(x-\mu-j\sigma^2 w)^2 - \frac{1}{2}\sigma^2 w^2 + jw\mu$$

$$-\frac{1}{2}(x-\mu-j\Sigma w)^T\Sigma^{-1}(x-\mu-j\Sigma w) - \frac{1}{2}w^T\Sigma w + jw^T\mu \xleftarrow{\text{n-Dim}} -\frac{1}{2\sigma^2}(x-\mu-j\sigma^2 w)^2 - \frac{1}{2}\sigma^2 w^2 + jw\mu$$

underbraces: $= 1$; "indep. to x." ; "remains"

$$\frac{1}{(2\pi)^{n/2}(\det\Sigma)^{1/2}}\int_{\mathbb{R}^n}\exp\left(-\frac{1}{2}(x-\mu-j\Sigma w)^T\Sigma^{-1}(x-\mu-j\Sigma w)\right)dx = 1$$

$$\therefore \quad \phi_X(w) = \exp\left(jw^T\mu - \frac{1}{2}w^T\Sigma w\right)$$

Linearity Property. $X \in \mathbb{R}^n, \quad X \sim N(\mu, \Sigma), \quad A \in \mathbb{R}^{m\times n}.$

Universal.

$$Y = AX \in \mathbb{R}^m \Rightarrow Y \sim N(A\mu, A\Sigma A^T)$$

$$\phi_Y(w) = E[\exp(jw^T Y)] = E[\exp(jw^T AX)] = E[\exp(j(A^Tw)^T X)] = \phi_X(A^T w) = \exp\left(j(A^Tw)^T\mu - \frac{1}{2}(A^Tw)^T\Sigma(A^Tw)\right)$$

$$= \exp\left(jw^T(A\mu) - \frac{1}{2}w^T(A\Sigma A^T)w\right) \Rightarrow \text{Gaussian Dist.}$$

$$\therefore \quad \mu_Y = A\mu, \quad \Sigma_Y = A\Sigma A^T, \quad Y \sim N(A\mu, A\Sigma A^T)$$

$$X = (X_1, \cdots, X_n)^T \sim N, \quad \widetilde{X} = (X_{n_1}, \cdots X_{n_k}), \quad (n_1 \cdots n_k) \subseteq (1, \cdots, n) \quad \widetilde{X} \sim N(\forall \text{ margin} \rightarrow \text{normal dist})$$

$$\widetilde{X} = \begin{bmatrix} X_{n_1} \\ \vdots \\ X_{n_k} \end{bmatrix} = \begin{bmatrix} \cdots & 1 & \cdots \\ \cdots & 1 & \cdots \\ \cdots & & 1 \cdots \end{bmatrix}\begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} \qquad \widetilde{X} \sim N(A\mu, A\Sigma A^T)$$

A ; n_k

Joint Gaussian $\rightleftarrows$ Boundary Gaussian

Example: $f(x_1, x_2) = \frac{1}{2\pi}\exp\left(-\frac{1}{2}(x_1^2+x_2^2)\right) + g(x_1,x_2)$

$$\begin{cases} \int_{\mathbb{R}} g(x_1, x_2)\,dx_1 = 0 \\ \int_{\mathbb{R}} g(x_1, x_2)\,dx_2 = 0 \end{cases} \qquad g(x_1, x_2) = \sin x_1 \sin x_2.$$

$$f'(x_1, x_2) = \frac{1}{2\pi}\exp\left(-\frac{1}{2}(x_1^2+x_2^2)\right)\left[1 + \sin x_1 \sin x_2\right] > 0$$

Boundary ✓, Joint ✗

Criterion for Joint Gaussian:

$$X \in \mathbb{R}^n, \quad X \sim N \iff \{ \forall \alpha \in \mathbb{R}^n, \; \alpha^T X \sim N \}$$

$$\phi_X(w) = E[\exp(jw^T X)] = \phi_{w^T X}(1) = \exp\left(j \cdot 1 \cdot \mu_{w^T X} - \frac{1}{2} \cdot 1 \cdot \sigma^2_{w^T X} \cdot 1\right) = \exp\left(j\mu_{w^T X} - \frac{1}{2}\sigma^2_{w^T X}\right)$$

$$\underbrace{w^T X \sim N}$$

$$\begin{cases} \mu_{w^T X} = E[w^T X] = w^T \mu_X \\ \sigma^2_{w^T X} = E[w^T X - w^T \mu_X]^2 \\ \qquad = E[w^T(X - \mu_X)]^2 \\ \qquad = w^T \Sigma_X \cdot w \end{cases}$$

$$\therefore \phi_X(w) = \exp(jw^T\mu_X - \frac{1}{2}w^T\Sigma_X w) \Rightarrow X \sim N(\mu_X, \Sigma_X)$$

$$X = (X_1, \cdots, X_n)^T \sim N, \quad E[X_i X_j] = EX_i \, EX_j \Rightarrow X_i, X_j \text{ independence.}$$

$$\Sigma_{ij} = E[X_i - EX_i)(X_j - EX_j) = E(X_i X_j) - EX_i EX_j = 0 \;\; (i \neq j) \Rightarrow \Sigma \text{ is diagonal.} \; (\Sigma = \text{diag.}(\sigma_1^2, \cdots, \sigma_n^2)$$

$$f_{X_1 \cdots X_n} = \frac{1}{(2\pi)^{n/2} \left(\prod_k \sigma_k^2\right)^{1/2}} \exp\left(-\frac{1}{2}\sum \frac{(x_k - \mu_k)^2}{\sigma_k^2}\right)$$

$$= \prod_k \left[\frac{1}{\sqrt{2\pi \sigma_k^2}} \exp\left(-\frac{1}{2}\frac{(x_k - \mu_k)^2}{\sigma_k^2}\right)\right] = \prod_k f_{X_k}(x_k)$$

In Gaussian.

Uncorrelated $\iff$ Independent.

PCA $\Rightarrow$ ICA (Independent)

$\downarrow$

Uncorrelated

Blind Source Separation

AIGC

Diffusion　MidJourney.

highly relevant to Stochastic Processes

$$X \sim N(0, I) \Rightarrow \widetilde{X} = \Sigma^{1/2}(X + \Sigma^{-1/2}\mu) \sim N(\mu, \Sigma)$$

$$X_0, X_1, X_2, \cdots, X_N \in \mathbb{R}^n \quad (X_k = \sqrt{1-\alpha_k}\, X_{k-1} + \sqrt{\alpha_k}\, \varepsilon_k, \quad \varepsilon_k \sim N(0, I), \text{ i.i.d.})$$

$$X_k \sim N! \rightarrow \text{Determine } X_m.!$$

$$\beta_k = 1 - \alpha_k, \quad X_k = \sqrt{\beta_k}\, X_{k-1} + \sqrt{1-\beta_k}\, \varepsilon_k = \sqrt{\beta_k}\,(\sqrt{\beta_{k-1}}\, X_{k-2} + \sqrt{1-\beta_{k-1}}\, \varepsilon_{k-1}) + \sqrt{1-\beta_k}\, \varepsilon_k$$

"Reparametric Trick"

$$= \sqrt{\beta_k \beta_{k-1}}\, X_{k-2} + \underbrace{\sqrt{\beta_k(1-\beta_{k-1})}\, \varepsilon_{k-1} + \sqrt{1-\beta_k}\, \varepsilon_k}_{\text{indep. Gaussian}}$$

$$\textcolor{blue}{N(0, \sqrt{\beta_k(1-\beta_{k-1})}\, I) + N(0, \sqrt{1-\beta_k}\, I) \Rightarrow N(0, \sqrt{1-\beta_k\beta_{k-1}}\, I)}$$

$$\therefore \begin{cases} X_k = \sqrt{\bar{\beta}}\, X_0 + \sqrt{1-\bar{\beta}}\, \varepsilon_0 \\ \bar{\beta} = \prod_k \beta_k \end{cases}$$

$$= \sqrt{\beta_k \beta_{k-1} \beta_{k-2}}\, X_{k-2} + \sqrt{1 - \beta_k \beta_{k-1}\beta_{k-2}}\, \varepsilon_0 = \cdots \cdots$$