Gaussian Processes 2

Recap.

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

Example.

$X_1, X_2, \cdots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$, $\bar{X} = \frac{1}{n}\sum_k X_k \Rightarrow \hat{\mu}$ $\quad \bar{S} = \frac{1}{n-1}\sum_k (X_k - \underbrace{\bar{X}}_{\text{contains } \bar{X}})^2 \Rightarrow \hat{\sigma}^2$

$\bar{X}$ and $\bar{S}$ are independent. given $X_k \sim N(\mu, \sigma^2)$

$$(n-1)\bar{S} = \sum_k (X_k - \bar{X})^2 = \sum_k (X_k^2 - 2\bar{X}X_k + \bar{X}^2) = \sum_k X_k^2 - 2\cdot\bar{X}\cdot n\bar{X} + n\bar{X}^2 = \sum_k X_k^2 - n\bar{X}^2$$

A (unitary mat.) doesn't change energy

Construct an unitary matrix

$$A = \begin{bmatrix} \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \cdots, \frac{1}{\sqrt{n}} \\ \\ * \\ \text{don't care} \end{bmatrix} \quad \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = A \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \Rightarrow$$

$Y_1 = \frac{1}{\sqrt{n}}\sum_k X_k$

$Y_1 = \sqrt{n}\bar{X}$, $\underbrace{\sum_i Y_i^2 = \sum_j X_j^2}$

$Y_1 \sim Y_n$ indep.

$\Sigma_Y = A\Sigma_X A^T = A(\sigma^2 I)A^T = \sigma^2(AIA^T) = \sigma^2 I$

$\therefore (n-1)\bar{S} = \sum_k X_k^2 - n\bar{X}^2 = \sum_k Y_k^2 - Y_1^2 = \sum_2^n Y_k^2$ indep. to $Y_1 = \sqrt{n}\bar{X}$

Degree of Freedom: $(n-1) \Rightarrow \bar{S}$ should be divided by $(n-1)$, not n

i.e. $\bar{S}$ indep. to $\bar{X}$ (works only in Gaussian Distribution) (Cochran)

why $\bar{S}$ has only $(n-1)$ DOF? Because $\bar{S} = \sum (X_k - \bar{X})^2$

Sample mean, not actual mean
1 DoF is subtracted here

Another Important Property:

$(X_1, X_2) \sim N\left(\begin{pmatrix}\mu_1 \\ \mu_2\end{pmatrix}, \begin{pmatrix}\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}\right)$ $\quad X_1, \mu_1 \in \mathbb{R}^m$, $X_2, \mu_2 \in \mathbb{R}^n$ $\quad$ Conditional Distribution? $(X_1 | X_2)$

$f_{X_2|X_1} = \frac{f_{X_1,X_2}}{f_{X_1}}$ $\Rightarrow$ Suppose to be a Gaussian Distribution , $\Sigma_{X_2|X_1} = ?$, $\mu_{X_2|X_1} = ?$

$$= C\exp\left(-\frac{1}{2}(X_1^T - \mu_1^T, X_2^T - \mu_2^T)\begin{bmatrix}\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22}\end{bmatrix}^{-1}\begin{pmatrix}X_1 - \mu_1 \\ X_2 - \mu_2\end{pmatrix} + \frac{1}{2}(X_1^T - \mu_1^T)\Sigma_{11}^{-1}(X_1 - \mu_1)\right)$$

Diagonalize

$$\begin{pmatrix}\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22}\end{pmatrix} \xrightarrow[\text{left}]{\begin{pmatrix} I & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I\end{pmatrix}\cdot\text{matrix}} \begin{pmatrix}\Sigma_{11} & \Sigma_{12} \\ 0 & -\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} + \Sigma_{22}\end{pmatrix} \xrightarrow[\text{right}]{\text{matrix}\cdot\begin{pmatrix} I & -\Sigma_{11}^{-1}\Sigma_{12} \\ 0 & I\end{pmatrix}} \begin{pmatrix}\Sigma_{11} & 0 \\ 0 & \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} I & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\begin{pmatrix} I & -\Sigma_{11}^{-1}\Sigma_{12} \\ 0 & I \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = \begin{pmatrix} I & 0 \\ \Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{pmatrix}\begin{pmatrix} I & \Sigma_{11}^{-1}\Sigma_{12} \\ 0 & I \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}^{-1} = \begin{pmatrix} I & -\Sigma_{11}^{-1}\Sigma_{12} \\ 0 & I \end{pmatrix}\begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & (\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1} \end{pmatrix}\begin{pmatrix} I & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix}$$

原式, $f_{X_2|X_1} = C\exp\left(-\frac{1}{2}(X_1^T-\mu_1^T, X_2^T-\mu_2^T)\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}^{-1}\begin{pmatrix} X_1-\mu_1 \\ X_2-\mu_2 \end{pmatrix} + \frac{1}{2}(X_1^T-\mu_1^T)\Sigma_{11}^{-1}(X_1-\mu_1)\right)$

$$= (X_1^T-\mu_1^T, X_2^T-\mu_2^T)\begin{pmatrix} I & -\Sigma_{11}^{-1}\Sigma_{12} \\ 0 & I \end{pmatrix}\begin{pmatrix} \Sigma_{11}^{-1} & \\ & (\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1} \end{pmatrix}\begin{pmatrix} I & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} X_1-\mu_1 \\ X_2-\mu_2 \end{pmatrix}$$

$$= \left(X_1^T-\mu_1^T, -(X_1^T-\mu_1^T)\Sigma_{11}^{-1}\Sigma_{12}+(X_2^T-\mu_2^T)\right)\begin{pmatrix} \Sigma_{11}^{-1} & \\ & (\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1} \end{pmatrix}\begin{pmatrix} X_1-\mu_1 \\ X_2-\mu_2-\Sigma_{21}\Sigma_{11}^{-1}(X_1-\mu_1) \end{pmatrix}$$

$$= (X_1^T-\mu_1^T)\Sigma_{11}^{-1}(X_1-\mu_1) + \left[(X_2^T-\mu_2^T)-(X_1^T-\mu_1^T)\Sigma_{11}^{-1}\Sigma_{12}\right](\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1}\left[X_2-\mu_2-\Sigma_{21}\Sigma_{11}^{-1}(X_1-\mu_1)\right]$$

$$\therefore f_{X_2|X_1} = C\exp\left(-\frac{1}{2}\left[X_2^T-\underbrace{\mu_2^T-(X_1^T-\mu_1^T)\Sigma_{11}^{-1}\Sigma_{12}}_{\text{mean}}\right]\underbrace{(\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})}_{\text{covariance matrix}}{}^{-1}\left[X_2-\mu_2-\Sigma_{21}\Sigma_{11}^{-1}(X_1-\mu_1)\right]\right)$$

$E[X_2|X_1]$ is a r.v. on X_1

pos. def.

$$\therefore X_2|X_1 \sim N(\mu_2+\Sigma_{21}\Sigma_{11}^{-1}(X_1-\mu_1), \Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$$

$X_1, X_2 \in \mathbb{R}^1$, $E[X_2|X_1] = \mu_2 + \frac{\sigma_{21}}{\sigma_{11}}(X_1-\mu_1)$ — linear adjustment. (Gaussian : Optimal !)

X_2 projection X_1 : prior knowledge

$\mathrm{var}(X_2|X_1) = \sigma_{22} - \frac{\sigma_{21}\sigma_{12}}{\sigma_{11}} = \sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}}$ (Randomness $\downarrow$: Gain Information (X_1) , Lose Uncertainty ($-\frac{\sigma_{12}^2}{\sigma_{11}}$)

X_1, X_2 uncorrelated : $\sigma_{12} = \sigma_{21} = 0 \Rightarrow \begin{cases} E[X_2|X_1] = \mu_2 = E[X_2] \\ \mathrm{var}(X_2|X_1) = \sigma_{22} = \mathrm{var}(X_2) \end{cases}$

Example : $X_1, X_2 \overset{i.i.d.}{\sim} N(0,1)$ $\quad E[X_1+X_2|X_1-X_2] = ?$, $E[(X_1+X_2)^2|X_1-X_2] = ?$

① $(X_1+X_2, X_1-X_2) \sim N(?,?)$

diagonal $\Rightarrow$ X_1', X_2' indep.

$E[X_1'|X_2'] = E[X_1'] = 0$.

$$\begin{pmatrix} X_1+X_2 \\ X_1-X_2 \end{pmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} I \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right) = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\right)$$

② $E[X_1'|X_2'] = E[X_1'] + \Sigma_{12}\Sigma_{22}^{-1}(X_2'-EX_2') = 0 + \frac{0}{2}(X_1-X_2-0) = 0$

X_1', X_2' indep. $E[X_1'^2|X_2'] = E[X_1'^2] = 2$

$E[X_2|X_1]$ has randomness, a r.v.

$(X_1,X_2)^T \sim N\left(\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}, \begin{bmatrix}\Sigma_{11}&\Sigma_{12}\\ \Sigma_{21}&\Sigma_{22}\end{bmatrix}\right)$, $X_2|X_1 \sim N(\mu_2+\Sigma_{21}\Sigma_{11}^{-1}(X_1-\mu_1), \Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$

$\Downarrow$

$Y_{X_2|X_1} \sim N(\cdots)$

$E[X_1^2|X_2'] = E[Y^2_{X_1'|X_2'}] = Var(Y_{X_1'|X_2'}) + E^2(Y_{X_1'|X_2'})$

Conditional Variable $\sim N(\cdots)$ conditional distribution.

$E[(X_1+X_2)^4|X_1-X_2] = ?$ $\quad Y_{X_1'|X_2'} = X_2+X_1|X_2-X_1 \sim N(0,2)$

$E[Y^4_{X_1'|X_2'}] \longrightarrow$ more general. $Y\sim N(0,\sigma^2)$, $E[Y^n] = \begin{cases}0, & n=2k-1\\ ? & n=2k\end{cases}$ $\quad \int y^n \exp(-\frac{y^2}{2\sigma^2})dy$

$$E(Y^{2k}) = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}} y^{2k}\exp(-\frac{y^2}{2\sigma^2})dy = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}} y^{2k-1}\, d(-\sigma^2\exp(-\frac{y^2}{2\sigma^2}))$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}}\Big[\underbrace{\int_{\mathbb{R}} d(-\sigma^2 y^{2k-1}\exp(-\frac{y^2}{2\sigma^2}))}_{0} - (-\sigma^2)(2k-1)\cdot\int_{\mathbb{R}} y^{2k-2}\exp(-\frac{y^2}{2\sigma^2})dy\Big]$$

$$= \sigma^2(2k-1)\cdot\frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb R} y^{2k-2}\exp(\frac{-y^2}{2\sigma^2})dy = \sigma^2(2k-1)\cdot E[Y^{2k-2}] = (\sigma^2)^{k-1}(2k-1)!!\cdot E[Y^2] = \sigma^{2k}\cdot(2k-1)!!$$

$$E[Y^n] = \begin{cases}0, & n=2k-1\\ (2k-1)!!\cdot\sigma^{2k}, & n=2k\end{cases}$$

$E[3X_1+2X_2|2X_1+3X_2] = ?$

$$\begin{bmatrix}3X_1+2X_2\\ 2X_1+3X_2\end{bmatrix} = \begin{bmatrix}3&2\\ 2&3\end{bmatrix}\begin{bmatrix}X_1\\ X_2\end{bmatrix} \sim N\left(\begin{bmatrix}3&2\\ 2&3\end{bmatrix}\begin{bmatrix}0\\ 0\end{bmatrix}, \begin{bmatrix}3&2\\ 2&3\end{bmatrix} I \begin{bmatrix}3&2\\ 2&3\end{bmatrix}\right) = N\left(\begin{bmatrix}0\\ 0\end{bmatrix}, \begin{bmatrix}13&12\\ 12&13\end{bmatrix}\right)$$

$$\underbrace{3X_1+2X_2|2X_1+3X_2}_{X_1'|X_2'} \sim N(\mu_1+\Sigma_{12}\Sigma_{22}^{-1}(X_2'-\mu_2'), \Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$$

$$= N(0+12\cdot\frac{1}{13}X_2', 13-\frac{144}{13})$$

$$= N(\frac{12}{13}(2X_1+3X_2), \frac{25}{13})$$

$\Rightarrow E[3X_1+2X_2|2X_1+3X_2] = \frac{12}{13}(2X_1+3X_2)$

$E[\cos(3X_1+2X_2)|2X_1+3X_2] = ?$

$Y_{1/2} = 3X_1+2X_2|2X_1+3X_2 \sim N(\frac{12}{13}(2X_1+3X_2), \frac{25}{13})$

$E[\cos Y_{1/2}] = E[\frac12(\exp(jY)+\exp(-jY))]$

$= \frac12\exp(-\frac{\frac{25}{13}}{2})\cos(\frac{12}{13}(2X_1+3X_2))$

$= \frac12\exp(-\frac{25}{26})\cos(\frac{12}{13}(2X_1+3X_2))$

$Y\sim N(\mu,\sigma^2)$, $E[\cos Y] = ?$

Use characteristic function:

$E[\cos Y] = E[\frac12(\exp(jY)+\exp(-jY))] = \frac12[\phi_Y(1)+\phi_Y(-1)]$

$\phi_Y = \exp(j\omega\mu - \frac{\omega^2\sigma^2}{2})$

$\Rightarrow E(\cos Y) = \exp(-\frac{\sigma^2}{2})\cdot\cos\mu$

$X \sim N(\mu_x, \Sigma_x)$, $Y = BX + z$, z, X independent, $z \sim N(0, \Sigma_z)$

1) $(X, z) \sim N$, $f_{X,z}(x,z) = f_X(x)\cdot f_z(z) = C\cdot \exp\left(-\frac{1}{2}(x-\mu_x)^T \Sigma_x^{-1}(x-\mu_x) - \frac{1}{2} z^T(\sigma^{-2} I) z\right)$

$$= C\cdot\exp\left(-\frac{1}{2}\left((x-\mu_x)^T, z^T\right)\begin{pmatrix}\Sigma_x^{-1} & \\ & \Sigma_z^{-1}\end{pmatrix}\begin{pmatrix}x-\mu_x\\ z\end{pmatrix}\right) \Rightarrow \text{Gaussian}$$

$$\begin{pmatrix}X\\ Y\end{pmatrix} = \begin{pmatrix}I & 0\\ B & I\end{pmatrix}\begin{pmatrix}X\\ z\end{pmatrix} \sim N$$

2) $X|Y \sim ?$ (X : state, z: noise, Y: observation)

$$X_2|X_1 \sim N\left(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1-\mu_1),\ \Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)$$

$$E[X|Y] = \mu + \Sigma_{XY}\Sigma_{YY}^{-1}(Y-\mu_Y)$$

X, z indep.

$$\Sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)^T] = E[(X-\mu_X)(BX+z-B\mu_X)^T] = \Sigma_X B^T + E(X-\mu_X)\overset{0}{E(z)} = \Sigma_X \cdot B^T$$

$$\Sigma_{YY} = E[(BX+z-B\mu_X)(BX+z-B\mu_X)^T] = B\Sigma_X B^T + \Sigma_z$$

$$\therefore E[X|Y] = \mu + \underbrace{\Sigma B^T}_{\Sigma_{XY}} \cdot (\underbrace{B\Sigma_X B^T+\Sigma_z}_{\Sigma_{YY}})^{-1}(Y - B\mu_X)$$ Bayesian.

$$\Sigma_{X|Y} = \Sigma_X - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX} = \Sigma_X - \Sigma_X B^T(B\Sigma_X B^T+\Sigma_z)^{-1} B\overset{\Sigma_X}{\Sigma_X^T}$$