清华大学电子工程系版权所有

第五章 高斯过程

5.1 Gauss 过程的基本定义

Gauss 过程之所以的连续参数连续状态随机过程，有两个重要原因。首先是许多实际问题涉及到大量的、微小的随机因素变量叠加后的结果，近似服从 Gauss 分布。其次，Gauss 过程容易解析计算为人们得到闭合表达式提供了便利。

定义 5.1 (Gauss 过程) 设有随机过程 $Z(t), t \in T$，如果 $\forall n, t_1, t_2, \ldots, t_n$，随机向量 $Z = (Z(t_1), Z(t_2), \ldots, Z_(t_n))^T$ 都服从 n 元 *Gauss* 分布，则称 $Z(t)$ 为高斯分布。

定义 5.2 (多元高斯分布) 若 n 元实随机向量 $Z = (Z(t_1), Z(t_2), \ldots, Z_(t_n))^T$ 的概率密度函数为

$$p(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \mathrm{e}^{-\frac{1}{2}(x-\mu)^{'}\Sigma^{-1}(x-\mu)}$$

其中 μ 为 n 维实相量，Σ 为 n 元对称正定矩阵，则称 Z 服从 n 元高斯分布。

- 非负性，$p(x) \geqslant 0$
- 正则性，$\int_{x\in\mathbb{R}^n} p_X(x)\mathrm{d}x = 1$
  因为 Σ 为正定，则存在非奇异矩阵 L, 使得 $\Sigma = L^{'}L$，做变换 $y = L^{'-1}(x-\mu)$ 后得到简单形式。
- $E[Z] = \mu, E[(z-\mu)(z-\mu)^{'}] \triangleq \mathrm{Cov}(Z) = \Sigma$
- 解读形式，如果一个 PDF 可以指数肩时的的 X 的一个二次型，那么这个 PDF 是多元高斯分布。通过 PDF 的解读可以判断是否是高斯分布。

定义 5.3 (多元 Gauss 分布的特征函数) 令 $t = (t_1, t_2, \ldots, t_n)^T$，则从 n 元随机向量 $Z = (Z_1, Z_2, \ldots, Z_n)^T$ 的特征函数为

$$\phi_Z(t) = E(\mathrm{e}^{jt^{'}Z}) = E(\mathrm{e}^{jt_iZ_i})$$

- 考虑 $AZ+b$ 的特征函数 $A \in R^{n\times n, b\in R^l}$, 则 $\phi_{AZ+b}(t) =$ 证明:

$$E[\mathrm{e}^{jt^{'}(AZ+b)}] = E[\mathrm{e}^{jt^{'}AZ}]\mathrm{e}^{jt^{'}b} = \phi_Z(A^{'}t)\mathrm{e}^{j}t^{'}b$$

- 混合矩. 若 $E[Z_1^{k_1}\ldots Z_n^{k_n}] < +\infty$, 则

$$E[Z_1^{k_1}\ldots Z_n^{k_n}] = (-j)^{k_1+k_2+\ldots+k_n}\frac{\partial\phi(t_1,\ldots,t_n)}{\partial t_1^{k_1}\ldots t_n^{k_n}}|_{t_1=\ldots=t_n=0}$$

清华大学电子工程系版权所有

证明：

$$\frac{\partial\phi(t_1,\ldots,t_n)}{\partial t_1}=\frac{\partial E[\mathrm{e}^{jt_1Z_1}\ldots\mathrm{e}^{jt_nZ_n}]}{\partial t_1}=E[\mathrm{e}^{jt_1Z_1}\ldots\mathrm{e}^{jt_nZ_n}\cdot jZ_1]$$

因此

$$\frac{\partial\phi(t_1,\ldots,t_n)}{\partial t_1}|_{t_1=t_2=\ldots=t_n=0}=jE[Z_1]$$

$$\begin{aligned}\frac{\partial\phi(t_1,\ldots,t_n)}{\partial t_1\partial t_2}&=\frac{\partial E[\mathrm{e}^{jt_1Z_1}\ldots\mathrm{e}^{jt_nZ_n}\cdot jZ_1]}{\partial t_2}\\&=E[\mathrm{e}^{jt_1Z_1}\ldots\mathrm{e}^{jt_nZ_n}\cdot jZ_1\cdot jZ_2]\end{aligned}$$

故

$$\frac{\partial\phi(t_1,\ldots,t_n)}{\partial t_1\partial t_2}|_{t_1=\ldots=t_n=0}=j^2E[Z_1Z_2]$$

- 设 $\phi(t_1,\ldots,t_n)$ 为 n 维向量 Z 的特征函数，$\phi_{Z_i}(t_i)$ 为 Z_i 的特征函数，$i=1,2,\ldots,n$，则 $Z_1,Z_2,\ldots,Z_n$ 相互独立的充要条件是

$$\phi(t_1\ldots t_n)=\phi_{Z_1}(t_1)\ldots\phi_{Z_n}(t_n)=\prod_{i=1}^{n}\phi_{Z_i}(t_i)$$

证明：随机变量的独立性可以通过它们的特征函数来刻画。
必要性：$Z_1,Z_2,\ldots,Z_n$ 相互独立，则 $\mathrm{e}^{jt_iZ_i}$ 也相互独立。

$$\phi(t_1,t_2,\ldots,t_n)=E[\mathrm{e}^{jt_1Z_1}\mathrm{e}^{jt_2Z_2}\ldots\mathrm{e}^{jt_1Z_1}]=E[\mathrm{e}^{jt_1Z_1}]E[\mathrm{e}^{jt_2Z_2}]\ldots E[\mathrm{e}^{jt_nZ_n}]=\phi_{Z_1}(t_1)\ldots\phi_{Z_n}(t_n)$$

充分性：设 $p_{Z_i}(X_i)$ 表示 Z_i 的密度函数，则 $\prod_{i=1}^{n}P_{Z_i}(x_i)$ 为 n 维分布密度。其特征函数为

$$\begin{aligned}&\int_{x_1\in R}\int_{x_n\in R}\prod_{i=1}^{n}P_{Z_i}(x_i)\prod_{i=1}^{n}\mathrm{e}^{jt_ix_i}dx_1\ldots dx_n\\&=\prod_{i=1}^{n}\int_{X_i\in R}P_{Z_i}(x_i)\mathrm{e}^{jt_ix_i}dx_i\\&=\prod_{i=1}^{n}\phi_{Z_i}(t_i)\\&=\phi(t_1,\ldots,t_n)\end{aligned}$$

- 多元高斯分布的特征函数

$$P_Z(x)\sim\mathcal{N}(\mu,\Sigma)\longrightarrow\phi_Z(t)=\mathrm{e}^{jt^{'}-\frac{1}{2}t^{'}\Sigma t}$$

证明：存在非奇异矩阵 L，使得 $\Sigma=LL^{'}$, 作线性变换，即

$$y=L^{'-1}(x-\mu),x-\mu=L^{'}y$$

清华大学电子工程系版权所有

$$\begin{aligned}
\phi_Z(t) &= \int_{x\in R^n} \mathrm{e}^{jt'x} P_Z(x)dx \\
&= \int_{y\in R^n} \mathrm{e}^{jt'(L'y+\mu)} \frac{1}{(2\pi)^{\frac{n}{2}}\Sigma^{\frac{1}{2}}} \mathrm{e}^{-\frac{1}{2}y'L'\Sigma^{-1}Ly} |\Sigma|^{\frac{1}{2}} dy \\
&= \mathrm{e}^{jt'\mu} \int_{y\in R^n} \mathrm{e}^{jt'L'y} \frac{1}{(2\pi)^{\frac{n}{2}}} \mathrm{e}^{-\frac{1}{2}y'y} dy \\
&= \mathrm{e}^{jt'\mu} \int_{y\in R^n} \mathrm{e}^{j\tau'y} \frac{1}{(2\pi^{\frac{n}{2}}} \mathrm{e}^{-\frac{1}{2}y'y} dy \\
&= \mathrm{e}^{jt'\mu} \prod_{i=1}^{n} \int_{y_i\in R} \mathrm{e}^{j\tau_i y_i} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}y_i^2} dy_i \\
&= \mathrm{e}^{jt'\mu} \prod_{i=1}^{n} \mathrm{e}^{-\frac{1}{2}\tau_i^2} \\
&= \mathrm{e}^{jt'\mu} \mathrm{e}^{-\frac{1}{2}\tau'\tau} \\
&= \mathrm{e}^{jt'\mu-\frac{1}{2}t'L'Lt} \\
&= \mathrm{e}^{jt'\mu-\frac{1}{2}t'\Sigma t}
\end{aligned}$$

定义 5.4 (多元 Gauss 分布利用特征函数的定义) 随机变量 $X \in \mathbb{R}$,若其特征函数为 $\phi_X(t) = \mathrm{e}^{jt'\mu-\frac{1}{2}t'\Sigma t}$，其中 $\mu \in \mathbb{R}$，Σ 非负定，则 $X \sim \mathcal{N}(\mu, \Sigma)$。

可见利用特征函数可以定义更广义的高斯分布。

- 高斯分布序列极限
  $X_n \xrightarrow{D} X$ 若 $\phi_{X_n(t)} \xrightarrow{D} \phi_X(t)$，利用特征函数知高斯分布的序列的极限仍为高斯分布，高斯过程的微分和积分仍是高斯过程。

- 例：$X_n \sim \mathcal{N}(a, \frac{1}{n})$，可得 $\phi_{X_n(t)} = \mathrm{e}^{jat-\frac{1}{2}\frac{t^2}{n}}$，当 $n \to \infty$，$\phi_{X_n(t)} = \mathrm{e}^{jat}, P(X = a) = 1.$

5.2 多元高斯分布的性质

多元 Gauss 分布具有许多其他分布不具备的良好性质，了解这些性质对于研究 Gauss 过程很重要。

- 边缘分布
  如果 $Z = Z_1, Z_2, \ldots, Z_n$ 服从 n 元 Gauss 分布，则它的任何一个子向量均服从 Gauss 分布。
  符号定义:n 元随机矢量 $Z_{1:n} \triangleq (Z_1, \ldots, Z_n)^T$, 即用下标集标示变量集。$A = \{a_1, \ldots, a_j\}$ 为 1 到 n 中 j 个不同数，$0 \leqslant j \leqslant n$, 则 Z_A 是 $Z_{1:n}$ 的 j 维子变量。$\Sigma_{A,B}$ 表示由 A 标示的行和 B 标示的列交叉组成的子矩阵。边缘分布的特征函数：

$$\phi_Z(t) = E[\mathrm{e}^{jt'Z}], t = t_{1:n} = (t_1, t_2, \ldots, t_n), Z = Z_{1:n} = (Z_1, Z_2, \ldots, Z_n)$$

  则子矢量的边缘分布

$$\phi_{Z_A}(t_A) = E[\mathrm{e}^{jt'_A Z_A}] = \phi_Z(t)|_{t_{\setminus A}=0}$$

清华大学电子工程系版权所有

特别的，若 $P_Z(x) \sim N(\mu, \Sigma)$ 则

$$\begin{aligned}
\phi_{Z_A}(t_A) &= E[\mathrm{e}^{jt'_A Z_A}] \\
&= E[\mathrm{e}^{jt' Z}]|_{t_{\setminus A}=0} \\
&= [\mathrm{e}^{jt'\mu - \frac{1}{2}t'\Sigma_t}]|_{t_{\setminus A}=0} \\
&= \mathrm{e}^{jt'_A\mu_A - \frac{1}{2}t'_A\Sigma_{AA}t_A}
\end{aligned}$$

知 Z_A 服从 Gauss 分布。

- 独立性

定理 5.1 *设 $Z = (Z_1, Z_2)^T$ 服从 Gauss 分布，均值向量为 $\mu_Z = (\mu_1, \mu_2)^T$，协方差矩阵为*

$$\Sigma_Z = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

则 Z_1, Z_2 相互统计独立的充要条件是 $\Sigma_{12} = 0$

证明：必要性是显然的，如果 Z_1 和 Z_2 相互统计独立，则

$$\Sigma_{12} = E[(Z_1 - \mu_1)(Z_2 - \mu_2)^T] = 0$$

充分性：当 $\Sigma_{12} = 0$ 时，协方差矩阵 Σ_Z 为分块对角阵

$$\Sigma_Z = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}$$

从而

$$\begin{aligned}
& p_Z(x_1, x_2) \\
&= \frac{1}{2\pi^{\frac{n}{2}} \begin{vmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{vmatrix}^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x_1 - \mu_1, x_2 - \mu_2)\begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}^{-1}(x_1 - \mu_1, x_2 - \mu_2)^T\right) \\
&= P_Z(x_1)P_Z(x_2)
\end{aligned}$$

推论 5.1 *服从 n 元 Gauss 分布的随机变量 $(Z_1, Z_2, \ldots, Z_n)$ 相互独立的充分必要条件是各元之间的协方差为 0.*

- 高阶矩

定理 5.2 *若 $Z = (Z_1, Z_2, Z_3, Z_4)^T$ 服从联合分布，且各分量的均值均为 0，则有*

$$E[Z_1 Z_2 Z_3 Z_4] = E(Z_1 Z_2)E(Z_3 Z_4) + E(Z_1 Z_3)E(Z_2 Z_4) + E(Z_1 Z_4)E(Z_2 Z_3)$$

- 线性变换

清华大学电子工程系版权所有

定理 5.3 设 n 维向量 $Z \sim \mathcal{N}(\mu, \Sigma)$, 则线性随机变换 $Y = AZ + b$ 仍然服从高斯分布 $N(A\mu + b, A\Sigma A^T)$

证明：

$$
\begin{aligned}
\phi_Y(t) &= \phi_Z(A^{'}t)e^{jt^{'}b} \\
&= \mathrm{e}^{j(A^{'}t)^{'}\mu - \frac{1}{2}(A^{'}t^{'})\sigma(A^{'}t)}e^{jt^{'}b} \\
&= \mathrm{e}^{jt^{'}(A\mu+b) - \frac{1}{2}t^{'}A\Sigma A^{'}t}
\end{aligned}
$$

即 $Y \sim \mathcal{N}(A\mu + b, A\Sigma A^T)$, 高斯分布的线性变换不变性，可以凭借这种不变性来判断一个随机变量是否服从多元高斯分布。

定理 5.4 $Z_{1:n}$ 服从 n 元高斯分布的充要条件是：任取 $C = (c_1, c_2, \ldots, c_n)^T \in R^n$, 线性组合 $C^{'}Z = \sum_{i=1}^{n} c_i Z_i$ 服从一元高斯分布。

证明：必要性是显然的，可以由线性变换不变性直接得到证明。
充分性：设 $Y = C^{'}Z$, 设 Z 的均值为 μ_Z, 协方差为 Σ_Z, 则 $E(Y) = C^{'}\mu_Z, Var(Y) = C^{'}\Sigma C$. 由于 Y 服从一元高斯分布，故

$$
\phi_Y(t) = E[\mathrm{e}^{jtY}] = E[\mathrm{e}^{jtC^{'}Z}] = e^{jC^{'}\mu_z t - \frac{1}{2}C^{'}\Sigma_Z C t^2}
$$

令 $t = 1$，则

$$
E[\mathrm{e}^{jC^{'}Z}] = \mathrm{e}^{jC^{'}\mu_z - \frac{1}{2}C^{'}\Sigma_Z C}
$$

由 C 的任意性可知 Z 服从 n 元联合高斯分布。

- 条件分布

  设 $X \sim \mathcal{N}(\mu, \Sigma)$, 其中，$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$,$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$,$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, 则

$$
\begin{aligned}
p(X_1|X_2) &= \mathcal{N}(\mu_{1|2}, \Sigma_{1|2}) \\
\mu_{1|2} &= \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_2) \\
\Sigma_{1|2} &= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}
\end{aligned}
$$

推导如下：
欲求 $p(X_1|X_2)$，考虑找到矩阵 C 使得 $Cov(Y, X_2) = 0$，其中

$$
\begin{pmatrix} Y \\ X_2 \end{pmatrix} = \begin{pmatrix} I & C \\ O & I \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, Y = X_1 + C \cdot X_2
$$

$$
p(X_1|X_2) = \frac{p(X_1, X_2)}{p(X_2)} = \frac{p(Y, X_2)}{p(X_2)} = p(Y)
$$

清华大学电子工程系版权所有

$$
\begin{aligned}
Cov(Y, X_2) &= E\left[(Y - EY)(X_2 - EX_2)^T\right] \\
&= E\left[(X_1 + CX_2 - \mu_1 - C\mu_2)(X_2 - \mu_2)^T\right] \\
&= \Sigma_{12} + C\Sigma_{22} = 0
\end{aligned}
$$

因此有$C = -\Sigma_{12}\Sigma_{22}^{-1}$,代入得

$$
\begin{aligned}
Y &= X_1 + CX_2 \\
EY &= \mu_1 + C\mu_2 \\
Cov(Y) &= E\left[(Y - EY)(Y - EY)^T\right] \\
&= E\left[(X_1 + CX_2 - \mu_1 - C\mu_2)(X_1 + CX_2 - \mu_1 - C\mu_2)^T\right] \\
&= \Sigma_{11} + C\Sigma_{22}C^T + C\Sigma_{21} + \Sigma_{12}C^T \\
&= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}
\end{aligned}
$$

5.3 PCA 简介

- 随机矢量的双正交展开

  由之前学过的卡洛定理我们知道，在区间 $[a, b]$ 上零均值，均方连续的随机过程 $X(t)$ 一定可以双正交展开：

  $$X(t) = \sum_{k=1}^{+\infty} X_k \phi_k(t)$$

  (1) $\phi_k(t)$ 特征函数

  $$\int_a^b C_X(t, s)\phi_k(s)\mathrm{d}s = \lambda_k \phi_k(t), t \in [a, b]$$

  (2) $X_k = \int_a^b X(t)\phi(t)\mathrm{d}t, E[X_i X_j] = \lambda_i \delta_{ij}$

  定理 5.5 对零均值的随机变量 $X \in \mathbf{R}$ 一定可以展开成

  $$X = \sum_{k=1}^{n} \xi_k e_k, e_k \in \mathbf{R}, k = 1, \ldots, n$$

  (1) e_k 是 $Cov(X) = \Sigma$ 的特征矢量，特征值 λ_k. $\|e_k\|^2 = 1$
  $e_i^T e_j = \delta_{ij}$

  (2) 同样有 $\xi = e_k^T X, E[\xi_i \xi_j] = \lambda_i \delta_{ij}$

  考虑对分布 $\mathcal{N}(\mu, \Sigma)$ 进行采样，可以利用正交分解，Σ 的特征矢量和特征值为 $e_1, \ldots, e_n$ 和 $\lambda_1, \ldots, \lambda_n$，可得样本 $X = \mu + \sum_{k=1}^n \xi_k e_k, \xi_k \sim \mathcal{N}(0, \lambda_k)$。

- PCA(Principle Component Analysis)

  PCA 在数据降维问题上得到了极为广泛的应用。这里我们不妨设零均值，对于随机矢量 X，目标是找到一个方向矢量 μ，使得

  $$\min_{u, \|\mu\|^2 = 1} E\left\|X - \mu^T X \mu\right\|^2$$

清华大学电子工程系版权所有

令 $\eta = \mu^T X$, 则

$$
\begin{aligned}
&E\left[(X-\eta\mu)^T(X-\eta\mu)\right] \\
&\quad = E[XX^T] - E[\eta X^T\mu] - E[\eta\mu^T X] + E[\eta\mu^T\mu\eta] \\
&\quad = E[XX^T] - E[\eta\eta]
\end{aligned}
$$

问题变为:

$$\max_{u,\|\mu\|^2=1} E[\eta^2] = E[\mu^T XX^T\mu] = \mu^T\Sigma_X\mu$$

利用拉格朗日乘子法

$$
\begin{gathered}
J(\mu,\alpha) = \mu^T\Sigma_X\mu - \alpha(\mu^T\mu - 1) \\
\frac{\partial J}{\partial\mu} = 2(\Sigma_X\mu - \alpha\mu) = 0
\end{gathered}
$$

即 μ 是 Σ_X 的特征矢量，主成分方向就是特征值最大的方向。

$$Var[\eta] = E[\eta]^2 = \mu^T\Sigma_X\mu = \mu^T\alpha\mu = \alpha\left\|\mu\right\|^2 = \alpha$$

可见第一主成分方向就是方差最大，X 投影其上变化最大的方向，一旦确定，则能被最大程度上确定。同样还能给出第二第三等主成分方向。

5.4 实高斯过程

定理 5.6 对于实高斯过程，严平稳 $\Leftrightarrow$ 宽平稳。
证明：设宽平稳 wss，均值常数为 μ_Z，则均方差函数仅与时间差相关，即

$$K(t_1,t_2) = K(t_1-t_2)$$

$\forall n,\forall t_1\ldots t_n,\forall h>0, Z(t_1)\ldots Z(t_n)$ 与 $Z(t_1+h)\ldots Z(t_n+h)$ 有相同的分布。
$Z(t_1)\ldots Z(t_n)$ 服从高斯分布，协方差为 $K(t_i-t_j), i,j=1,\ldots,n$
$Z(t_1+h)\ldots Z(t_n+h)$ 亦服从高斯分布，协方差为 $K(t_i+h-t_j-h), i,j=1,\ldots,n$

定理 5.7 对 $GP\begin{pmatrix}Z(t)\\Y(t)\end{pmatrix}$ $Z(t),Y(t)$ 独立 $\Leftrightarrow Z(t)$ 与 $Y(t)$ 不相关.

证明：$\forall n,m,t_1\ldots t_n,\tau_1\ldots\tau_m$，则

$$P_{Z(t_1)\ldots Z(t_n),Y(\tau_1)\ldots Y(\tau_m)}(x_1\ldots,x_n,y_1\ldots y_m) = P_{Z(t_1)\ldots Z(t_n)}(x_1\ldots x_n)P_{Y(\tau_1)\ldots Y(\tau_m)}(y_1\ldots y_m)$$

$$\forall t_1,t_2, C_{Z,Y}(t_1,t_2) = 0$$

此时 $(Z(t_1)\ldots Z(t_n)\,Y(\tau_1)\ldots Y(\tau_m))'$ 服从高斯分布，其协方差阵为分块对角。

定理 5.8 高斯矢量序列的均方极限服从高斯分布。

定理 5.9 均方收敛 $\Rightarrow$ 依概率收敛。
证明：若 $Z_n \stackrel{m.s.}{\to} Z$，根据 $Markov$ 不等式

$$\forall\varepsilon>0, P(|Z_n-Z|^2\geqslant\varepsilon^2)\leqslant\frac{E|Z_n^{'}-Z|^2}{\varepsilon^2}$$

清华大学电子工程系版权所有

等式左边：$E|Z_n^{'}-Z|^2 \to 0(n\to\infty)$
等式右边：$\forall\varepsilon, P(|Z_n-Z|\geqslant\varepsilon)\to 0(n\to\infty)$

定理 5.10 若高斯过程 $Z(t), t\in T$ 均方可导，则 $Z^{'}(t)$ 也是高斯过程。

定理 5.11 若高斯过程 $Z(t), t\in T$ 在 T 上均方可积，$Y(t)\triangleq Z(\tau)h(t,\tau)d\tau, t\in T$ 也是高斯过程。

定理 5.12 $Z(t), t\in T$ 为 $GP \Leftrightarrow \forall h(t)$,

$$\int_{t\in T} h^T(t)Z(t)dt$$

服从高斯分布。
证明：必要性显然。
充分性：欲证 $\forall n, \forall t_1\dots t_n\in T, Z(t_1)\dots Z(t_n)$ 服从高斯分布。考虑 $\forall C_i$ 与 $Z(t)$ 同维，

$$h(t)=\sum_{i=1}^n C_i\int(t-t_i)$$

则

$$\int_{t\in T} h(t)^T Z(t)dt = C^{'}(Z(t_1)\dots Z(t_n))^T$$

服从高斯分布，由 C 的任意性可知 $Z(t_1)\dots Z(t_n)$ 服从高斯分布。

定理 5.13 高斯过程 $Z(t), t\in T$ 通过线性系统 $h(t,\tau)$ 后输出 $Y(t), t\in T$，则 $\begin{pmatrix} Z(t)\\ Y(t)\end{pmatrix}$ 为高斯过程。

5.5 窄带 Gauss 过程

5.5.1 rayleigh 分布和 Rician 分布

定义 5.5 (Rician 分布) 设 Z, Y 相互独立,服从联合高斯分布,$E(Z)=\mu_1, E(Y)=\mu_2, Var(Z)=Var(X)=\sigma^2$, 则

$$p_{ZY}=\frac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{1}{2\sigma^2}[(x-\mu_1)^2+(y-\mu_2)^2]}$$

设 $\mu_1=A\cos\phi, \mu_2=A\sin\phi$

$$p(x,y)=\frac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{1}{2\sigma^2}(x^2+y^2+A^2-2Ax\cos\phi-2Ay\sin\phi)}$$

极坐标变换：$x=r\cos\theta, y=r\sin\theta$

$$J=\begin{pmatrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial\theta}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial\theta}\end{pmatrix}=\begin{pmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{pmatrix}$$

$$p(r,\theta)=p(x,y)|\frac{\partial(x,y)}{\partial(r,\theta)}|=\frac{r}{2\pi\sigma^2}e^{-\frac{1}{2\sigma^2}[r^2+A^2-2Ar\cos(\theta-\phi)]}$$

清华大学电子工程系版权所有

边缘分布：

$$p(r) = \frac{r}{\sigma^2}\mathrm{e}^{-\frac{r^2+A^2}{2\sigma^2}}\frac{1}{2\pi}\int_0^{2\pi}\mathrm{e}^{\frac{Ar}{\sigma^2}\cos(\theta-\phi)}d\theta$$
$$= \frac{r}{\sigma^2}\mathrm{e}^{-\frac{r^2+A^2}{2\sigma^2}}I_0(\frac{Ar}{\sigma^2}), r \geqslant 0$$

这里 $I_0(Z)$ 是修正的 *Bessel* 函数，定义如下

$$I_0(Z) = \frac{1}{2\pi}\int_0^{2\pi}\mathrm{e}^{z\cos\theta}d\theta$$

$A = 0$ 是 *Rayleign* 分布；$A \gg 1$，在 A 附近近似高斯分布。

定义 5.6 (Rayleigh 分布) 如果 $A = 0$，则

$$p(r,\theta) = \frac{r}{2\pi\sigma^2}\mathrm{e}^{-\frac{r^2}{2\sigma^2}}, r \geqslant 0, \theta \in [0, 2\pi]$$

对 r 积分得 $p(\theta) = \frac{1}{2\pi}, \theta \in [0, 2\pi]$，进而

$$p(r) = \frac{r}{\sigma^2}\mathrm{e}^{-\frac{r^2}{2\sigma^2}}, r \geqslant 0$$

$$E(R) = \sqrt{\frac{\pi}{2}}\sigma, Var(R) = \frac{4-\pi}{2}\sigma^2$$

此时 R 与 Θ 相互统计独立。

5.5.2 零均值窄带高斯过程

- $$\begin{pmatrix} Z_C(t) \\ Z_S(t) \end{pmatrix} = \begin{pmatrix} \cos w_C(t), \sin w_C(t) \\ -\sin w_C(t), \cos w_C(t) \end{pmatrix}\begin{pmatrix} Z(t) \\ \hat{Z(t)} \end{pmatrix}$$

  故 $Z_C(t), Z_S(t)$ 是联合高斯过程。

- $$R_{Z_C}(0) = R_{Z_S}(0) = R_Z(0) = R_{\hat{Z}}(0) = \sigma_Z^2$$

  $$R_{Z_C Z_S}(0) = -R_{Z_C Z_S}(0) = 0$$

  即同一时刻 $Z_C(t), Z_S(t)$ 不相关 $\longrightarrow$ 独立

- 包络过程：$R(t) = \sqrt{Z_C(t)^2 + Z_S(t)^2}$
  相位过程：$\Theta(t) = \arctan\frac{Z_S(t)}{Z_C(t)}$

- $\begin{pmatrix} R(t) \\ \Theta(t) \end{pmatrix}$ 联合宽平稳。

清华大学电子工程系版权所有

5.6 高斯线性回归

考虑利用 M 组数据 $x^{(i)} = (x_1^{(i)}, \dots, x_n^{(i)})^T$ 和权值 $w = (w_1, \dots, w_n)^T$ 对目标 $y^{(i)}$ 进行拟合，记

$$X = (x^{(1)}, \dots, x^{(M)}), Y = (y^{(1)}, \dots, y^{(M)})^T$$

- 方法 A: 线性回归
  利用最小二乘法，$Y = X^T w$, 求解即伪逆：
  $$w = (XX^T)^{-1} XY$$

- 方法 B: 高斯回归
  $$\begin{aligned}
  y^{(i)} &= f(x^{(i)}) + e^{(i)}, \quad e^{(i)} \sim \mathcal{N}(0, \sigma^2) \\
  &= {x^{(i)}}^T w + e^{(i)}, \quad w \sim \mathcal{N}(0, \Sigma_p) \\
  Y &= X^T w + E\text{，经过增广} \\
  \begin{pmatrix} Y \\ w \end{pmatrix} &= \begin{pmatrix} I & X^T \\ 0 & I \end{pmatrix} \begin{pmatrix} E \\ w \end{pmatrix} \\
  \begin{pmatrix} Y \\ w \end{pmatrix} &\sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 I + X^T \Sigma_p X & X^T \Sigma_p \\ \Sigma_p X & \Sigma_p \end{pmatrix} \right) \\
  w|Y &\sim \mathcal{N}\left(\Sigma_p X (\sigma^2 I + X^T \Sigma_p X)^{-1} Y, \Sigma_p - \Sigma_p X (\sigma^2 I + X^T \Sigma_p X)^{-1} X^T \Sigma_p \right) \\
  &= \mathcal{N}(\mu_w, \Sigma_w)
  \end{aligned}$$
  因此给一个新的数据 x^*
  $${x^*}^T w \sim \mathcal{N}({x^*}^T \mu_w, {x^*}^T \Sigma_w x^*)$$

- 方法 C: 高斯过程观点
  定义高斯过程 $\{f(x), x \in \mathbb{R}^n\}$，均值为 $Ef(x) = E[X^T w] = 0$
  $$C(u, v) = E\left[f(u) f(v)^T\right] = e[u^T w w^T v] = u^T \Sigma_p v$$
  可以得到有限维高斯分布：
  $$\begin{pmatrix} y^{(1)} \\ \dots \\ y^{(M)} \\ f(x^*) \end{pmatrix} = \begin{pmatrix} f(x^{(1)}) + e^{(1)} \\ \dots \\ f(x^{(M)}) + e^{(M)} \\ f(x^*) \end{pmatrix}$$
  其均值为 0，协方差矩阵为
  $$\begin{pmatrix} C(x^{(i)}, x^{(j)}) + \sigma^2 \mathbf{I}_{M*M} & C(x^{(i)}, x^{(*)}) \\ C(x^{(*)}, x^{(j)}) & C(x^{(*)}, x^{(*)}) \end{pmatrix}$$
  则有 $f(x^*)|y^{(1)}, \dots, y^{(M)}$ 均值为 ${x^*}^T \Sigma_p X (\sigma^2 I + X^T \Sigma_p X)^{-1} Y$
  利用核的思想，直接考虑协方差 $C(u, v)$
  如：$C(u, v) = k \exp(-\frac{\|u-v\|^2}{2s^2})$