高斯过程回归方法及其应用

Gaussian processes regression and its application

XX
XXX

目录

Definition 1 (高斯分布)

一个连续型随机变量 X 其概率密度函数 (probability density function, pdf) 若为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad x \in \mathbb{R}. \tag{0.1}$$

其中 $\mu \in \mathbb{R}$, $\sigma^2 \geqslant 0$, 则称 X 服从均值为 μ 方差为 σ^2 的高斯分布 (Gaussian distribution) 或正态分布 (normal distribution). 记作

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

高斯过程

随机变量的概念很容易推广到多维随机变量 (multivariate random variables) 或称为随机向量 (random vectors).

$$Y = (Y_1, \ldots, Y_d)^T$$

Y 每一个分量都是定义在**同一个概率空间**下的随机变量. 随机向量 Y 的一个样本 y 应有 $y \in \mathbb{R}^d$.

$$\mathrm{E}(Y) = (\mathrm{E}(Y_1), \ldots, \mathrm{E}(Y_d))^T. \tag{0.2}$$

$$\mathrm{var}(Y) = \mathrm{E}(YY^T) - \mathrm{E}(Y)\mathrm{E}(Y^T) = \mathrm{E}[(Y - \mathrm{E}(Y))(Y - \mathrm{E}(Y))^T]. \tag{0.3}$$

$$\mathrm{cov}(Y, Z) = \mathrm{E}(YZ^T) - \mathrm{E}(Y)\mathrm{E}(Z^T). \tag{0.4}$$

Definition 2 (多维高斯分布)

若有随机向量 $Y=(Y_1,\ldots,Y_d)^T$, 其分量的任意线性组合而成的**随机变量**都服从于高斯分布, 即

$$\forall\alpha\in\mathbb{R}^d,\ \alpha_1Y_1+\cdots+\alpha_dY_d=\alpha^TY\sim\mathcal{N}. \tag{0.5}$$

此时称随机向量 Y 服从于 d 维高斯分布, Y 也被称为高斯随机向量.

d 维高斯分布的概率密度函数由下式给出.

$$f_Y(y)=\frac{1}{|2\pi\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(y-\mu)^T\Sigma^{-1}(y-\mu)\right)\qquad y\in\mathbb{R}^d. \tag{0.6}$$

可简记为,

$$Y\sim f(y)=\mathcal{N}(\mu,\Sigma) \tag{0.7}$$

高斯过程

$$f_Y(y) = \frac{1}{|2\pi\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(y-\mu)^T\Sigma^{-1}(y-\mu)\right) \qquad y \in \mathbb{R}^d.$$

$$Y \sim f(y) = \mathcal{N}(\mu, \Sigma)$$

其中 μ 是 d 维的均值向量 (mean vector), 其每个分量 $\mu_i = \mathrm{E}(Y_i)$, 而 Σ 是一个 $d \times d$ 的矩阵, 其 i 行 j 列的分量 $\Sigma_{i,j} = \mathrm{cov}(Y_i, Y_j)$, 矩阵 Σ 被称为协方差矩阵 (covariance matrix).

多维高斯分布的特性由均值向量 μ 与协方差矩阵 Σ 唯一确定. 可以证明但此处不证, 协方差矩阵总是对称半正定 (symmetric and positive semi-definite) 的, 即,

$$\Sigma_{i,j} = \Sigma_{j,i}. \tag{0.8}$$

$$\forall \alpha \in \mathbb{R}^d,\ \alpha^T\Sigma\alpha \geqslant 0. \tag{0.9}$$

性质 0.1
对于高斯向量来说, 各分量之间不相关等价于各分量之间独立.

$$\Sigma = \begin{pmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_d^2 \end{pmatrix} \tag{0.10}$$

$$\begin{aligned} f(y) &= \frac{1}{|2\pi\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(y-\mu)^T\Sigma^{-1}(y-\mu)\right) \\ &= \frac{1}{\sqrt{2\pi\sigma_1^2} \times \cdots \times \sqrt{2\pi\sigma_d^2}} \exp\left(-\sum_{i=1}^{d} \frac{(y_i-\mu_i)^2}{2\sigma_i^2}\right) \\ &= \prod_{i=1}^{d} f(y_i). \end{aligned} \tag{0.11}$$

高斯过程

$$Y \sim f(y) = \mathcal{N}(\mu, \Sigma)$$

为了研究高斯分布的条件分布, 我们把高斯随机向量划分为两部分, $Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$, Y_1 与 Y_2 都可能为高斯随机向量或只是单变量 (univariate). 同时也对应将均值向量与协方差矩阵分块,

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \tag{0.12}$$

$$\Sigma = \begin{pmatrix} \Sigma_{1,1} & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{pmatrix} \tag{0.13}$$

性质 0.2
已知 Y_2 的条件下 Y_1 的条件分布也是高斯分布. 若 Y_1 为随机向量, 则该条件分布是多维高斯分布, 且该多维高斯分布的均值向量与协方差函数由下式给出,

$$Y_1|Y_2 \sim \mathcal{N}(\mu_c, \Sigma_c).$$

$$\mu_c = \mu_1 + \Sigma_{1,2}\Sigma_{2,2}^{-1}(Y_2 - \mu_2). \tag{0.14}$$

$$\Sigma_c = \Sigma_{1,1} - \Sigma_{1,2}\Sigma_{2,2}^{-1}\Sigma_{2,1}. \tag{0.15}$$

其中 $Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$, $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$, $\Sigma = \begin{pmatrix} \Sigma_{1,1} & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{pmatrix}$.

分块矩阵求逆定理 (inversion of a partitioned matrix). 设矩阵 A 为分块矩阵, 其逆阵为 A^{-1},

$$A = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}, \qquad A^{-1} = \begin{pmatrix} \tilde{P} & \tilde{Q} \\ \tilde{R} & \tilde{S} \end{pmatrix}. \tag{0.16}$$

其中 P 与 $\tilde{P}$ 为 $n_1 \times n_1$ 的方阵, S 与 $\tilde{S}$ 为 $n_2 \times n_2$ 的方阵. 在已知 P,Q,R,S 的条件下, 可由下式计算得到 $\tilde{P}$,$\tilde{Q}$,$\tilde{R}$,$\tilde{S}$.

$$\begin{cases} \tilde{P} = N. \\ \tilde{Q} = -NQS^{-1}. \\ \tilde{R} = -S^{-1}RN. \\ \tilde{S} = S^{-1} + S^{-1}RNQS^{-1}. \end{cases} \tag{0.17}$$

其中 $N = (P - QS^{-1}R)^{-1}$.

证明: 不失一般性, 我们假定 Y 是经过中心化的, 即 $\mu = 0$. 对于高斯分布我们通常只需关注指数部分, 从多维高斯分布的 pdf 出发, 我们得到,

$$\begin{aligned} f_Y(y) &\propto \exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right) \\ &\propto \exp\left(-\begin{pmatrix} y_1^T & y_2^T \end{pmatrix}\begin{pmatrix} \Sigma_{1,1} & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{pmatrix}^{-1}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right). \end{aligned} \tag{0.18}$$

下面求分块矩阵的逆, 令,

$$\begin{pmatrix} \Sigma_{1,1} & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{pmatrix}^{-1} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{0.19}$$

高斯过程

由分块矩阵求逆定理及协方差矩阵性质 $\Sigma_{1,2} = \Sigma_{2,1}^T$ 可求得,

$$\begin{cases} A = N. \\ B = -N\Sigma_{1,2}\Sigma_{2,2}^{-1}. \\ C = -\Sigma_{2,2}^{-1}\Sigma_{2,1}N. \\ D = \Sigma_{2,2}^{-1} + \Sigma_{2,2}^{-1}\Sigma_{2,1}N\Sigma_{1,2}\Sigma_{2,2}^{-1}. \end{cases} \tag{0.20}$$

上式中 $N = (\Sigma_{1,1} - \Sigma_{1,2}\Sigma_{2,2}^{-1}\Sigma_{2,1})^{-1}$.

$$\begin{aligned} f_Y(y) &\propto \exp\left(-\begin{pmatrix} y_1^T & y_2^T \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) \\ &\propto \exp(-y_1^TAy_1 - y_2^TCy_1 - y_1^TBy_2 - y_2^TDy_2) \\ &\propto \dots \\ &\propto \exp\left(-(y_1 - \Sigma_{1,2}\Sigma_{2,2}^{-1}y_2)^T(\Sigma_{1,1} - \Sigma_{2,2}^{-1}\Sigma_{2,1})^{-1}(y_1 - \Sigma_{1,2}\Sigma_{2,2}^{-1}y_2)\right). \end{aligned} \tag{0.21}$$

高斯过程

$$f_Y(y) \propto \exp\left(-(y_1 - \underbrace{\Sigma_{1,2}\Sigma_{2,2}^{-1}y_2}_{\mu_c})^T(\underbrace{\Sigma_{1,1} - \Sigma_{2,2}^{-1}\Sigma_{2,1}}_{\Sigma_c})^{-1}(y_1 - \underbrace{\Sigma_{1,2}\Sigma_{2,2}^{-1}y_2}_{\mu_c})\right).$$

在 y_2 已知的条件下, $\mu_c = \Sigma_{1,2}\Sigma_{2,2}^{-1}y_2$ 与 $\Sigma_c = \Sigma_{1,1} - \Sigma_{2,2}^{-1}\Sigma_{2,1}$ 是确定的, 故 $Y_1|Y_2$ 服从均值为 μ_c, 协方差矩阵为 Σ_c 的高斯分布, 即

$$Y_1|Y_2 \sim \mathcal{N}(\mu_c, \Sigma_c).$$

如果不考虑进行中心化, 则只需将 y_1 与 y_2 分别用 $y_1 - \mu_1$ 与 $y_2 - \mu_2$ 替代, 即得到结果,

$$\mu_c = \mu_1 + \Sigma_{1,2}\Sigma_{2,2}^{-1}(Y_2 - \mu_2).$$
$$\Sigma_c = \Sigma_{1,1} - \Sigma_{1,2}\Sigma_{2,2}^{-1}\Sigma_{2,1}.$$

$\square$

多维随机向量的概念可以推广到随机过程 (stochastic processes).

Definition 3 (高斯过程)

若一个定义在域 $D \in \mathbb{R}^d$ 上的随机过程 Z, 对于 $\forall n \in \mathbb{N},\ \forall x_i \in D,\ (Z(x_1), \ldots, Z(x_n))$ 是一个高斯随机向量, 则 Z 是高斯过程 (Gaussian Process, GP).

与高斯分布类似, 高斯过程的性质可以由定义在 D 上的均值函数 (mean function) 与定义在 $D \times D$ 上的协方差函数 (covariance function) 完全确定. 通常用 $m(x)$ 与 $k(x, x')$ 来表示均值函数与协方差函数.

$$m(x) = \mathrm{E}(Z(x)) \tag{0.22}$$

$$k(x, x') = \mathrm{cov}(Z(x), Z(x')) \tag{0.23}$$

于是高斯过程 Z 可简记为

$$Z \sim \mathcal{GP}\left(m(x), k(x, x')\right)$$

高斯过程

在实际应用中总是先将数据预处理成是 0 均值的, 即使得 $m(x) = 0$. 这样不仅在理论推导时带来便利, 在实际运算中也提高了效率.

对于高斯过程的一个样本 $(Z(x_1), \ldots, Z(x_n))$ 是服从多维高斯分布的, 这个高斯分布的协方差矩阵 K 由协方差函数计算得到

$$K = \begin{pmatrix} k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \cdots & k(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n, x_1) & k(x_n, x_2) & \cdots & k(x_n, x_n) \end{pmatrix} \tag{0.24}$$

协方差函数中存在着最终确定高斯过程的参数, 这些参数也被称为超参数 (hyper-parameters).

高斯过程

协方差函数也被称为核函数 (kernel function). 核函数形式繁多也有多种分类方法, 常见协方差函数如下,

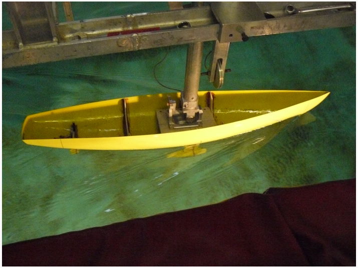
表 1 常见协方差函数

核函数	表达式	参数
Constant	σ_f^2	σ_f^2
Linear	$\sum_{i=1}^{d}\sigma_i^2 x_i x_i'$	$\{\sigma_1^2,\ldots,\sigma_d^2\}$
Polynomial	$(xx'+\sigma_f^2)^p$	σ_f^2
Squared Exponential	$\sigma_f^2\exp\left(-\frac{\lVert x-x'\rVert^2}{2l^2}\right)$	σ_f^2, l
Matérn	$\frac{2^{1-v}}{\Gamma(v)}\left(\frac{\sqrt{2v}r}{l}\right)^v K_v\left(\frac{\sqrt{2v}r}{l}\right)$	v, l
Exponential	$\sigma_f^2\exp\left(-\frac{\lVert x-x'\rVert}{l}\right)$	σ_f^2, l
Rational Quadratic	$\sigma_f^2\left(1+\frac{\lVert x-x'\rVert^2}{2\alpha l^2}\right)^{-\alpha}$	σ_f^2, α, l

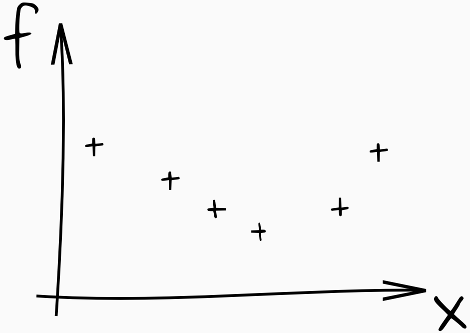
高斯过程回归模型 - 统计模型优劣

为什么使用统计模型 (statistical modelling)?

- 样本量小
  - 样本不易采集 (如现实世界的实验).
  - 样本获取的成本太高 (如破坏性测试, 耗时的数值仿真).
  - ...

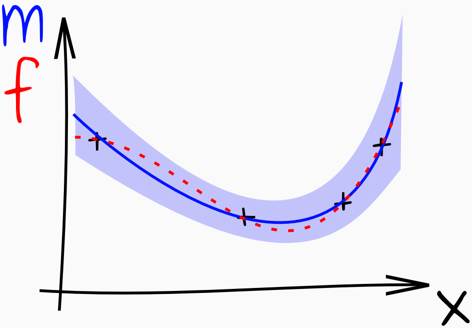


- 无法事先确定模型的形式.



高斯过程回归模型 - 统计模型优劣

- 统计模型使用数据得到预测分布进行预测，同时可给出模型误差置信区间，量化模型预测的不确定性.



模型假设

$$y = f(\boldsymbol{x}) + \varepsilon.$$

假设域 $D \in \mathbb{R}^d$, $\boldsymbol{x} \in D$, $y \in \mathbb{R}$, $\varepsilon \sim \mathcal{N}(0,\ \sigma_n^2)$, 并设 f 是定义在 D 上的高斯过程.

$$f \sim \mathcal{GP}\left(0, k(\boldsymbol{x},\ \boldsymbol{x}')\right).$$

经过采样, 得到训练数据集 $\{(\boldsymbol{x}_i, y_i) | i = 1, \ldots, n\}$, 记 $\{(\boldsymbol{X}, \boldsymbol{y})\}$. 同样可获得测试数据集 $\{(\boldsymbol{X}_*, \boldsymbol{y}_*)\}$.

由高斯过程理论, 可知联合分布服从高斯分布.

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{f}_* \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(\boldsymbol{X},\boldsymbol{X}) + \sigma_n^2 I & K(\boldsymbol{X},\boldsymbol{X}_*) \\ K(\boldsymbol{X}_*,\boldsymbol{X}) & K(\boldsymbol{X}_*,\boldsymbol{X}_*) \end{bmatrix}\right).$$

此处 $K(\boldsymbol{X},\boldsymbol{X}')$ 为简记, 当 $\boldsymbol{X} = (\boldsymbol{x}_1 \ldots \boldsymbol{x}_m)$ 与 $\boldsymbol{X}' = (\boldsymbol{x}'_1 \ldots \boldsymbol{x}'_d)$ 时意义为下式,

$$K(\boldsymbol{X},\boldsymbol{X}') = \begin{bmatrix} k(\boldsymbol{x}_1,\boldsymbol{x}'_1) & \ldots & k(\boldsymbol{x}_1,\boldsymbol{x}'_d) \\ \vdots & \ddots & \vdots \\ k(\boldsymbol{x}_m,\boldsymbol{x}'_1) & \ldots & k(\boldsymbol{x}_m,\boldsymbol{x}'_d) \end{bmatrix}_{m\times d}.$$

$$y = f(\boldsymbol{x}) + \varepsilon.$$

考虑噪声方差.

$$\operatorname{cov}(y_p, y_q) = k(\boldsymbol{x}_p, \boldsymbol{x}_q) + \sigma_n^2 \delta_{pq}. \tag{0.25}$$

其中 δ_{pq} 为克罗内克函数 (Kronecker delta function), 其形式为

$$\delta_{pq} = \begin{cases} 1 & p = q, \\ 0 & p \neq q. \end{cases} \tag{0.26}$$

从协方差矩阵的角度看,

$$\operatorname{cov}(\boldsymbol{y}) = K(\boldsymbol{X}, \boldsymbol{X}) + \sigma_n^2 I. \tag{0.27}$$

高斯过程回归模型 - 建立模型

由高斯分布条件分布的性质, 高斯过程回归的预测分布为

$$\boldsymbol{f}_* | \boldsymbol{X}_*, \boldsymbol{X}, \boldsymbol{y} \sim \mathcal{N}(\bar{\boldsymbol{f}}_*,\ \mathrm{cov}(\boldsymbol{f}_*))$$

$$\boxed{\begin{aligned} \bar{\boldsymbol{f}}_* &= K\big(\boldsymbol{X}_*, \boldsymbol{X}\big)[K(\boldsymbol{X}, \boldsymbol{X}) + \sigma_n^2 I]^{-1} \boldsymbol{y} \\ \mathrm{cov}(\boldsymbol{f}_*) &= K(\boldsymbol{X}_*, \boldsymbol{X}_*) - K(\boldsymbol{X}_*, \boldsymbol{X})[K(\boldsymbol{X}, \boldsymbol{X}) + \sigma_n^2 I]^{-1} K(\boldsymbol{X}, \boldsymbol{X}_*) \end{aligned}}$$

高斯过程回归模型 - 模型求解

可用最大似然估计方法对模型内的参数进行估计.

最大似然估计法本质上是使用对数损失函数 (logarithmic loss function) 的经验风险最小化 (empirical risk minimization, ERM) 方法. 核心的思想就是在已有输入与结果的条件下, 调整模型参数使这样的输入输出关系发生的概率最大.

在高斯过程模型下, 设模型参数向量为 $\boldsymbol{\theta}$, 模型参数主要来自于协方差函数 $k(\boldsymbol{x}, \boldsymbol{x}')$ 与噪声方差 σ_n^2.

假设各次采样值 $\boldsymbol{y} \in \mathbb{R}^d$独立同分布于同一个高斯分布. 记 $\Sigma_{xx} = K(\boldsymbol{X}, \boldsymbol{X})$, $\Sigma_{yy} = \mathrm{cov}(\boldsymbol{y})$.

$$\begin{aligned} p(\boldsymbol{y}; \boldsymbol{\theta}) &= \frac{1}{\sqrt{(2\pi)^d |\Sigma_{yy}|}} \exp\left(-\frac{1}{2}\boldsymbol{y}^T \Sigma_{yy}^{-1} \boldsymbol{y}\right) \\ &= \frac{1}{\sqrt{(2\pi)^d |\Sigma_{yy}|}} \exp\left(-\frac{1}{2}\boldsymbol{y}^T (\Sigma_{xx} + \sigma_n^2 I)^{-1} \boldsymbol{y}\right). \end{aligned} \tag{0.28}$$

n 次采样可写出似然函数

$$L(\boldsymbol{Y};\boldsymbol{\theta}) = \prod_{i=1}^{n} p(\boldsymbol{y}_i;\boldsymbol{\theta}) = \left(\frac{1}{\sqrt{(2\pi)^d|\Sigma_{yy}|}}\right)^n \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\boldsymbol{y}_i^T\Sigma_{yy}^{-1}\boldsymbol{y}_i\right). \quad (0.29)$$

$$\log(L(\boldsymbol{Y};\boldsymbol{\theta})) = -\frac{dn}{2}\log(2\pi) - \frac{n}{2}\log(|\Sigma_{yy}|) - \frac{1}{2}\sum_{i=1}^{n}\boldsymbol{y}_i^T\Sigma_{yy}^{-1}\boldsymbol{y}_i. \quad (0.30)$$

为求解 $\max_{\boldsymbol{\theta}} \ L(\boldsymbol{Y};\boldsymbol{\theta})$.

$$\min_{\boldsymbol{\theta}} \ -2\log(L(\boldsymbol{Y};\boldsymbol{\theta})) = \min_{\boldsymbol{\theta}} \ dn\log(2\pi) + n\log(|\Sigma_{yy}|) + \sum_{i=1}^{n}(\boldsymbol{y}_i^T\Sigma_{yy}^{-1}\boldsymbol{y}_i).$$

$$\boldsymbol{\theta} = \operatorname{argmin} dn\log(2\pi) + n\log(|\Sigma_{yy}|) + \sum_{i=1}^{n}(\boldsymbol{y}_i^T\Sigma_{yy}^{-1}\boldsymbol{y}_i). \quad (0.31)$$

高斯过程回归模型 - 模型求解

应用 GPR 建模的一般过程

采集样本数据

$\downarrow$

确定先验模型, 选择核函数, 确定超参数初始值

$\downarrow$

使用训练样本对模型进行训练, 优化目标似然函数

$\downarrow$

得到最优的超参数集, 最终确定后验模型

$\downarrow$

计算 $\bar{\boldsymbol{f}}_*$ 与 $\mathrm{cov}(\boldsymbol{f}_*)$

高斯过程回归模型应用 - 仿真

高斯过程回归输入为一维的例子 (150/200)

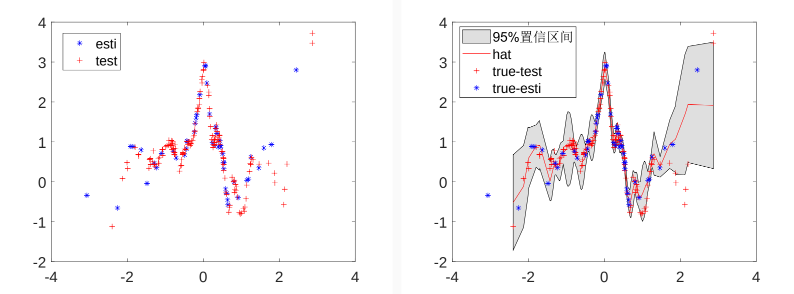


图 1　仿真

工具: Matlab + GPML ToolBox

```
meanfunc = {@meanSum, {@meanLinear,
@meanConst}};

covfunc = {@covMaterniso, 3};
```

高斯过程回归模型应用 - 大气 CO_2 浓度预测

大气 CO_2 浓度预测取是《$GPML$》中的一个例子, 作者选取了 1958 年到 2003 年以月为间隔的 545 个数据 (大气平均 CO_2 浓度) 用于估计模型参数, 设计的模型协方差函数如下,

$$k_1(x,x') = \theta_1^2 \exp\left(-\frac{(x-x')^2}{2\theta_2^2}\right)$$

$$k_2(x,x') = \theta_3^2 \exp\left(-\frac{(x-x')^2}{2\theta_4^2} - \frac{2sin^2(\pi(x-x'))}{\theta_5^2}\right)$$

$$k_3(x,x') = \theta_6^2 \left(1 + \frac{(x-x')^2}{2\theta_8\theta_7^2}\right)^{-\theta_8}$$

$$k_4(x_p,x_q) = \theta_9^2 \exp\left(-\frac{(x_p-x_q)^2}{2\theta_{10}^2}\right) + \theta_{11}\delta_{pq}$$

$$k(x,x') = k_1(x,x') + k_2(x,x') + k_3(x,x') + k_4(x,x')$$

高斯过程回归模型应用 - 大气 CO_2 浓度预测

预测结果，

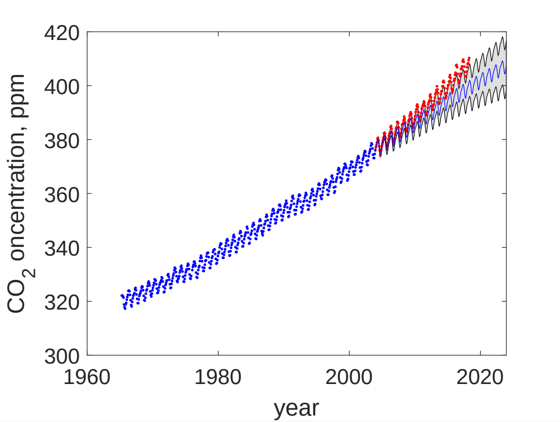


图 2 CO_2 浓度预测

参考 i

张江帆.
基于高斯过程回归的锂电池数据处理.
北京交通大学, 2017.

C. E. Rasmussen, C. K. I. Williams.
Gaussian processes in machine learning.
the MIT Press, Cambridge, 2006.

Carl Edward Rasmussen, Hannes Nickisch.
The gpml toolbox version 4.1.
2017.

N. Durrande.
Lecture notes on gaussian process regression, 2017.

...

Questions?

Thank you!