

第一次 project

一. 主题介绍

- 本次大作业的主题是——马氏链蒙特卡洛方法
- 马氏链蒙特卡洛 (Markov Chain Monte Carlo, MCMC) 方法, 是马氏链理论的一个重要应用。从 1950 年萌芽, 马氏链蒙特卡洛方法在实践中不断发展, 逐渐成长为一个颇具分量的理论分支, 广泛应用于各种学科领域 (如信息科学、物理、化学、生物学、金融、材料等) 的科学计算, 展示出越来越强大的威力。
- MCMC 的一个简短介绍见文献 [1]——林元烈编著的《应用随机过程》第 3.5 节 P^n 的极限性态与平稳分布。英文的简短介绍可见文献 [2]——《Pattern Recognition and Machine Learning》Chapter 11 Sampling Methods。更系统的介绍见文献 [3]。

二. 作业题目

(a) 二维高斯分布的估计。

用 Metropolis-Hastings 算法, 对下述二维高斯分布进行随机采样

$$\mathcal{N}\left\{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \middle| \begin{pmatrix} 5 \\ 10 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}\right\}$$

- 利用采样结果估计二维高斯分布的相关系数 ρ 。
- 改变建议分布的选取进行仿真比较, 可以从拒绝概率, 收敛速度等角度对算法进行分析。

(b) 基于 MCMC 的径向基函数 (Radial Basis Function, RBF) 网络模型选择方法。

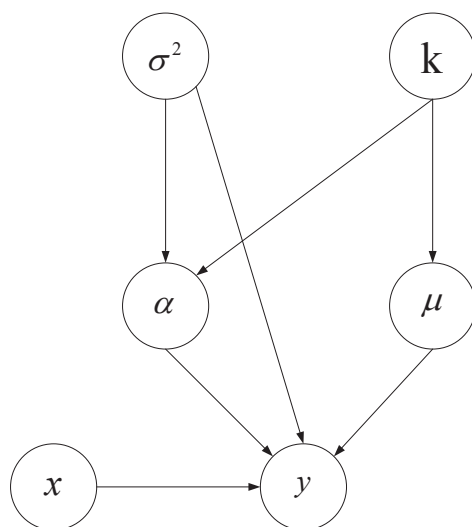
RBF 网络是一种经典的神经网络结构。由于 RBF 函数能够逼近任意的非线性函数, RBF 网络通常被应用于处理系统内的难以解析的规律性, 具有良好的泛化能力, 并有很快的学习收敛速度, 已成功应用于非线性函数逼近、时间序列分析、数据分类、模式识别、信息处理、图像处理、系统建模、控制和故障诊断等。RBF 模型可以简单表示为以下形式 [4]:

$$\mathbf{y}_{N \times c} = \mathbf{D}(\mu_{k \times d}, \mathbf{x}_{N \times d}) \alpha_{(1+d+k) \times c} + \mathbf{n}_{N \times c}$$

\mathbf{y} 为观测样本, N 为样本点数, c 为每个样本点的维度; \mathbf{x} 为 N 个观测样本点对应的变量, d 为每个变量的维度; 矩阵 \mathbf{D} 有如下形式

$$\begin{pmatrix} 1 & \mathbf{x}_{1,1} & \dots & \mathbf{x}_{1,d} & \phi(\mathbf{x}_1, \mu_1) & \dots & \phi(\mathbf{x}_1, \mu_k) \\ 1 & \mathbf{x}_{2,1} & \dots & \mathbf{x}_{2,d} & \phi(\mathbf{x}_2, \mu_1) & \dots & \phi(\mathbf{x}_2, \mu_k) \\ \vdots & \vdots & & & & & \vdots \\ 1 & \mathbf{x}_{N,1} & \dots & \mathbf{x}_{N,d} & \phi(\mathbf{x}_N, \mu_1) & \dots & \phi(\mathbf{x}_N, \mu_k) \end{pmatrix}$$

$\phi(\mathbf{x}, \mu)$ 表示 RBF 函数, k 为 RBF 函数的个数, μ 为函数的中心; α 为加权系数, \mathbf{n} 是服从高斯分布的白噪声, $\mathbf{n}_i \sim \mathcal{N}(\mathbf{0}, \sigma_i^2), i = 1, 2, \dots, c$ 。RBF 模型的参数集合可用 $\mathcal{M}_s = (\mu, \alpha, \sigma^2, \mathbf{k})$ 表示, 下图简单说明了 RBF 模型的基本结构:



经典的 RBF 模型选择方法有很多种，如 AIC 方法 [5]、MDL 方法 [6]、BIC 方法 [7] 以及 MAP 方法 [8]。文献 [4] 的作者提出了一种基于 RJ(Reversible Jump)-MCMC[9] 及模拟退火算法 (Simulated Annealing algorithm, SA) 的 RBF 模型选择方法，可以对 RBF 网络进行更好的学习。更多与 RBF 模型选择相关的资料请参见文献 [10, 11]。

- i. 设计和实现一个运用 RJ-MCMC 的实验，表明你理解了 RJ-MCMC。
- ii. 采用 [4] 中提出的方法利用给定的数据训练并测试 RBF 模型，与上述 AIC 等经典方法选择几种进行比较，分析不同方法的性能。
- iii. 请大家自行调研相关文献，尝试采用本文档中未提到的其他方法解决 RBF 模型选择问题。

附件文件说明：附件中包含两个数据文件 “data1.mat” 和 “data2.mat”，是两个用于 RBF 模型训练和测试的数据集，具体说明如下：

- (a) “data1.mat” 文件中包括 x 、 y 、 x_{test} 三个变量，其中 x 、 y 为训练 RBF 模型需要用到的数据， x 是 1000×2 的矩阵，是模型的输入变量； y 是 1000×2 的矩阵，是模型的输出结果， x_{test} 为 200×2 的测试数据
- (b) “data2.mat” 文件中包括 x 、 y 、 x_{test} 三个变量， x 是 1000×3 的矩阵， y 是 1000×2 的矩阵， x_{test} 为 200×3 的测试数据，定义同上。

对于每个训练好的模型，将对应的测试数据 x_{test} 输入模型后，将会得到一个 200×2 维的输出矩阵，请将使用同样方法得到的两个输出矩阵存在同一个.mat 文件中，命名为“方法学号.mat”进行提交。详细要求见“具体要求”部分。

三. 具体要求

- (a) 希望同学充分调研和阅读相关文献，积极动脑 + 动手，取得有自己见解的结果，整理成最终报告。
- (b) **最终提交包括：**

i. 报告

报告的书写要求参见《Project 报告撰写建议》。

- ii. 源程序 务必包含: 自己所有的原始程序、所有方法的输出数据: 请将使用 RJMCMC 方法训练得到的 RBF 模型输出的两个 200×2 维矩阵存放在“RJMCMC 学号.mat”中, 如“RJMCMC2015xxxxxx.mat”, 用“data1.mat”训练得到的模型对应的输出矩阵命名为“v1”, 另一个输出矩阵命名为“v2”。如果你还使用了其他方法对 RBF 模型进行了训练, 请将对应模型的输出结果存放在“方法学号.mat”中, 例如你采用了 AIC 方法训练得到了 RBF 模型并进行了测试, 应将输出结果存放在“AIC2015xxxxxx.mat”中, 两个矩阵同样命名为“v1”和“v2”。

将以上两项一起压缩打包, 命名为“学号 _ 姓名.rar”进行提交。

- (c) 评分标准: 报告书写清晰和规范, 工作新意及深入程度, 工作量及完整程度, 模型的输出结果与原始数据的贴近程度。
- (d) 一旦发现抄袭, 计零分。
- (e) 请大家在规定截止时间前提交。晚交的处理方法如下: 按晚交天数, 以 90% 的几何级数进行折扣。晚交时间在 (0, 24 小时], 按 90% 折扣。晚交时间在 (24 小时, 48 小时], 按 90%*90% 折扣。以此类推。

参考文献

- [1] 林元烈, 应用随机过程. 清华大学出版社, 2002.
- [2] C. M. Bishop, *Pattern recognition and machine learning*. springer, 2006.
- [3] J. S. Liu, *Monte Carlo strategies in scientific computing*. Springer Science & Business Media, 2008.
- [4] C. Andrieu, N. D. Freitas, and A. Doucet, “Reversible jump mcmc simulated annealing for neural networks,” in *Sixteenth Conference on Uncertainty in Artificial Intelligence*, pp. 11–18, 2000.
- [5] H. Akaike, “A new look at the statistical model identification,” *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.
- [6] G. Schwarz, “Estimating the dimension of a model,” *Annals of Statistics*, vol. 6, no. 2, pp. págs. 15–18, 1978.
- [7] J. Rissanen, “Stochastic complexity,” *Journal of the Royal Statistical Society*, vol. 49, no. 3, pp. 223–239, 1987.
- [8] P. M. Djuric, “Asymptotic map criteria for model selection,” *IEEE Transactions on Signal Processing*, vol. 46, no. 10, pp. 2726–2735, 2002.
- [9] P. J. Green, “Reversible jump markov chain monte carlo computation and bayesian model determination,” *Biometrika*, vol. 82, no. 4, pp. 711–732, 1995.
- [10] C. Andrieu, J. D. Freitas, and A. Doucet, “Robust full bayesian learning for neural networks,” 1999.
- [11] C. Andrieu, P. M. Djurić, and A. Doucet, “Model selection by mcmc computation.,” *Signal Processing*, vol. 81, no. 1, pp. 19–37, 2001.