

分布式存储与计算 第三次作业

提交时间：2022 年 11 月 15 日 18:30

1. 使用逆累积分布函数法产生密度函数为 $f(x) = \frac{3x^2}{2}, -1 \leq x \leq 1$ 的随机数。试用 Scala 编写程序产生 10000 个上述分布的随机数。
2. 假设随机向量 $(X, Y)^T$ 来自于二元正态分布：

$$N\left(\begin{pmatrix} 10.0 \\ -10.0 \end{pmatrix}, \begin{pmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{pmatrix}\right)$$

随机产生 10^5 个这样的向量, 请使用 MapReduce 的方法计算下列期望所对应的样本值：

- a) $E(X^2Y^2)$
 - b) $E\begin{pmatrix} X^2 & XY \\ XY & Y^2 \end{pmatrix}$
 - c) $E\begin{pmatrix} \sin(X^2) & \cos(XY) \\ \cos(XY) & \sin(Y^2) \end{pmatrix}$
3. 首先从混合二元正态分布
$$0.4 \times N\left(\begin{pmatrix} 1.0 \\ 2.0 \end{pmatrix}, \begin{pmatrix} 2.0 & 1.0 \\ 1.0 & 2.0 \end{pmatrix}\right) + 0.6 \times N\left(\begin{pmatrix} 5.0 \\ 6.0 \end{pmatrix}, \begin{pmatrix} 4.0 & 1.0 \\ 1.0 & 4.0 \end{pmatrix}\right)$$
中产生 10000 个随机向量。
 4. 使用上题中产生的数据, 编写分布式 EM 算法 (使用 4 个计算节点) 来估计相关参数。