

## Mathematics check points

1. 是否能从本质上理解下式？

$$p(A) = p(A, B) + p(A, \text{not } B) = p(A|B) \times p(B) + p(A|\text{not } B) \times p(\text{not } B)$$

2. 下列两式说明什么？

$$\begin{aligned} p(A) &= p(A, B) \\ p(A, B) &= p(A) \times p(B) \end{aligned}$$

3. Bayes Rule 的理解与使用

4. 为什么方差的定义为  $\sigma^2 = \sum_i p(x_i)(x_i - \mu)^2$ ，但我们用下式去估计方差？

$$\frac{1}{N-1} \sum_{i=1}^N (x_i - \hat{\mu})^2$$

5. 两个随机变量相关系数为 0 意味着什么？
6. 极大似然估计 MLE 的原理与方法
7. 极大后验概率 MAP 的原理与方法
8. 向量、矩阵的基本运算
9. 向量的 Lp 范数 Vector Lp-Norm 的定义与性质
10. 两个向量正交的定义与性质，如何理解正交？ (orthogonal vectors)
11. 正交矩阵是什么，它有什么性质？ (orthogonal matrices)
12. 半正定矩阵的定义是什么，我们如何简单理解半正定矩阵，它有什么性质？  
(positive semi-definite matrices)
13. 我们为什么需要在机器学习中了解优化理论？
14. 一个优化问题是什么样的？
15. 什么是一个多元函数的梯度和 Hessian 矩阵？
16. 什么是局部最小值，什么是整体最大值？
17. 什么是凸集 (convex set)，可以举一个例子吗？
18. 什么是凸函数 (convex function)，可以举一个例子吗？
19. 凸函数的 first order 和 second order 有什么性质？
20. 梯度下降方法的原理与做法？
21. 梯度下降中，步长的设定需要考虑什么？
22. 如何对一个优化问题写出它的拉格朗日函数 (Lagrangian)？
23. 如何通过拉格朗日函数得出一个优化问题的对偶问题？
24. 什么样的优化问题是一个凸问题？
25. 如果一个优化问题是凸问题，那么求解原问题和求解它的对偶问题，会产生一样的结果吗？