

21188142: 课程综合实践 II (数据要素市场)

2024-2025 学年短学期

HW 1: 数据定价基础

教师: 刘金飞, 助教: 吴一航

日期: 2024 年 7 月 3 日

1.1 偏好的性质

使用偏好关系 \succeq 的完备性和传递性, 以及偏好关系 \sim 的定义证明:

如果 $x \sim y$ 且 $y \sim z$, 则 $x \sim z$.

证明: 由于 $x \sim y$, 则 $x \succeq y$ 且 $y \succeq x$; 由于 $y \sim z$, 则 $y \succeq z$ 且 $z \succeq y$. 由于 \succeq 是传递的, 所以 $x \succeq z$ 且 $z \succeq x$, 即 $x \sim z$. ■

1.2 有预算约束的效用最大化

已知 Frank 用于购买土豆和牛肉的预算约束为 p 美元, 土豆和牛肉的市场价格分别为 p_1 和 p_2 美元. 当 Frank 购买 x_1 盎司土豆和 x_2 盎司牛肉时, 他的效用函数为 $U(x_1, x_2) = x_1 x_2^2$. 求 Frank 对土豆和牛肉的需求函数.

解: 本题是上课例子的特例, 代入 $\alpha = 1, \beta = 2$ 即可得到答案为 $x_1 = \frac{p}{3p_1}, x_2 = \frac{2p}{3p_2}$.

1.3 无套利原则

1. 判断以下情况是否存在套利机会, 如果是, 请构造一个套利策略; 如果不是, 请说明理由.

(a) 在外汇市场中, 你预期欧元会下跌, 即未来的 1 欧元换成人民币可能比现在的要便宜;
不存在; 注意套利的定义要求无风险, 欧元下跌只是个人预期, 不一定会发生, 存在风险

(b) 在 PPT 的查询例子中, 如果令

$Q_2 = \text{select gender, count(*) from User group by gender}$

且 $p(Q_1, D) = 3, p(Q_2, D) = 2, p(Q_3, D) = 4, p(Q_4, D) = 1$;

存在; (以下两点回答一点即可) Q_1 的价格高于 Q_2 违反了信息无套利, 可以通过购买 Q_2 的信息然后出售 Q_1 的信息来获得套利; Q_3 的价格大于 Q_2 和 Q_4 的价格之和违反了组合无套利, 可以通过购买 Q_2 和 Q_4 的信息然后出售 Q_3 的信息来获得套利

- (c) 机器学习模型市场中，如果训练任务是一个线性回归模型，版本化的方法是对其中某个系数添加高斯噪声 $\varepsilon \sim N(0, \sigma^2)$ ，每次添加的噪声之间相互独立。并且添加 $\varepsilon_1 \sim N(0, 2)$ 的模型价格为 10，添加 $\varepsilon_2 \sim N(0, 3)$ 的模型价格为 6，添加 $\varepsilon_3 \sim N(0, 4)$ 的模型价格为 3。

存在；首先显然噪声方差越小说明模型与准确值越接近，因此方差越小价格越高是合理的；其次因为我们是估计参数，例如 $y = ax + b$ 的 a ，那么我们可以购买两个加噪的 a 取平均获得噪声更小的参数：（以下情况回答一点即可）例如购买两个添加了 ε_3 的模型，那么 $\frac{\varepsilon_3 + \varepsilon_3}{2} \sim N(0, 2)$ ，只需要 6 元即可获得价格为 10 的模型；也可以组合 $\frac{\varepsilon_2 + \varepsilon_3}{2} \sim N(0, 1.75)$ ，只需要 9 元即可获得比价格为 10 的模型更好的模型

2. 在查询定价无套利等价条件的讨论中，我们要求集合函数 $f : 2^S \rightarrow \mathbb{R}^+$ 满足单调性和次可加性，请判断以下函数是否符合要求。

(a) $f(A) = |A|, \forall A \subseteq S$;

符合；单调性显然，次可加性根据容斥原理显然。

(b) $f(A) = \max_{a \in A} a, \forall A \subseteq S$.

不符合次可加性；例如 $A = \{-1\}, B = \{1\}$ ，则 $f(A \cup B) = 1 > 0 = f(A) + f(B)$ 。