

EC '22: ACM Conference on Economics and Computation

Preference Dynamics Under Personalized Recommendations

——建模在个性化推荐的场景下人偏好（或观点）的演化过程

自动化所, 蔡怀广, 202228014628021

目录

CONTENTS

01. 背景

02. 建模

03. 结论

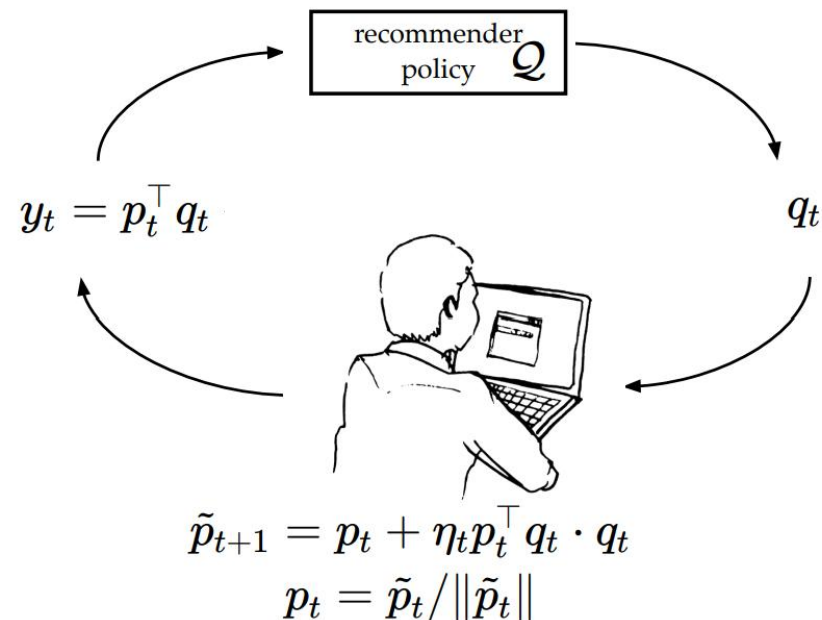
04. 评价

背景

- 信息爆炸的时代，推荐系统成为解决信息过载的主流方案
- 有智能手机的用户几乎不可能接触不到推荐系统
- 推荐系统的存在影响了人的观点：
 - 回声室效应：在网络空间内，人们经常接触相对同质化的人群和信息，听到相似的评论，倾向于将其当作真相和真理，不知不觉中窄化自己的眼界和理解，走向故步自封甚至偏执极化。
 - 过滤气泡效应：我们看到的内容经过了算法（基于平台对用户喜好的观察）的过滤。我们不喜欢或不同意的新闻内容会被自动过滤掉，而这会缩小我们的认知范围。
- 推荐系统的存在对人群观点的影响？
 - 大家的观点是趋向还是分散？
 - 推荐系统有没有加深人的偏见，让人变得激进？从而人群走向对立？

建模推荐系统对个体偏好的影响

- **Biased Assimilation** proposed by HJMR [2019]
- 原文: preferences become more aligned with content that is enjoyed, and anti-aligned with content that is disliked.
- 解释: 如果 t 时刻用户 p_t 看到的推荐内容 q_t 是他喜欢的 ($p_t^T q_t > 0$), 那么接下来用户的偏好 p_{t+1} 会变得和内容 q_t 更加一致; 反之, 如果 t 时刻用户 p_t 看到的推荐内容 q_t 是他讨厌的 ($p_t^T q_t < 0$), 那么接下来用户的偏好 p_{t+1} 会变得更加远离内容 q_t .
- 目标, 用Regret来衡量: $R(T) = \sum_{t=0}^{T-1} r_\star - r_t$
 1. 最大化用户满意程度: $r_t = p_t^T q_t$, 对应着用户购买物品的概率或者用户浏览视频的时间。
 2. 最大化用户兴趣不改变程度: $r_t = p_0^T p_t$, 对应着推荐系统要尽可能降低对用户的影响, 不要去操纵用户的观点。



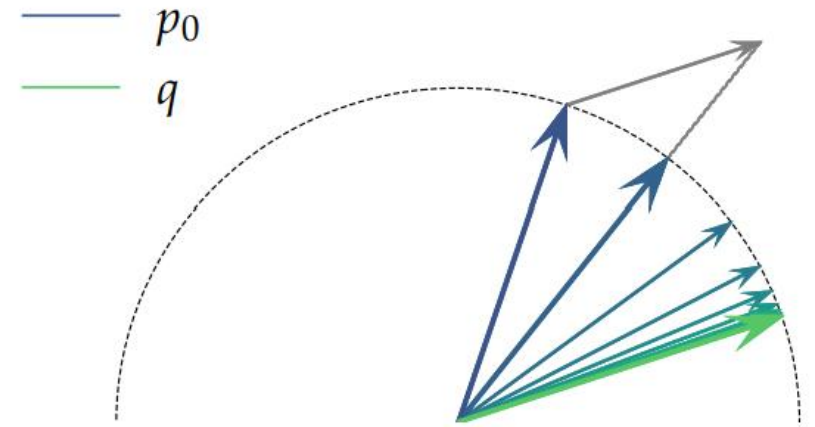
符号:

- t 时刻用户的偏好记为 p_t
- t 时刻展示给用户的内容记为 q_t
- 用户对内容的喜爱程度记为 $p_t^T q_t$

$$\tilde{p}_{t+1} = p_t + \eta_t p_t^\top q_t \cdot q_t$$

结论1：最大化用户满意度是很简单的

- 简单的策略就能让用户的满意程度达到最优
 1. 找到用户不讨厌的一个内容 $p_0^\top q > 0$
 2. 一直给用户推荐这个内容 $q_t = q$
- 算法有效的原因：用户的偏好 p_t 会向喜欢的内容 q 方向移动，所以第一步只要找到用户不讨厌的一个内容（用户点赞了某个内容或者看了某个视频很长时间），接下来一直推荐这个内容，用户的偏好 p_t 最终会变得和 q 一模一样，此时用户的满意程度就是1，和最佳策略一样。



PROPOSITION 3. *If $q_t = q$ for all $t \geq 0$ and $p_0^\top q > 0$, then the cumulative regret is at most constant:*

$$R(T) \leq C_Y ((p_0^\top q)^{-2} - 1), \quad C_Y = \begin{cases} \frac{(\eta^2+1)}{\eta^2+2\eta} & \eta_t = \eta \\ \frac{s^2 \pi^2}{6} & \eta_t = \frac{\eta}{t+s} \end{cases}$$

$$\tilde{p}_{t+1} = p_t + \eta_t p_t^\top q_t \cdot q_t$$

结论1续：给所有用户都推荐相同内容会导致观点极化

- 类似于上面的结果，如果平台的策略是给所有人都推荐相同内容 ($q_t = q$)。
- 初始时，有一些用户喜欢 q ，最终这些用户的偏好都变为 q ；
- 初始时，有一些用户不喜欢 q ，最终这些用户的偏好都变为 $-q$ ；
- 于是人群观点发生极化：一派人的观点是 q ，另一派人的观点是 $-q$ ，不存在中间可能！
- 解释生活中的例子：假设每天在新闻上都报道美国枪击案 (q)，一部分人最终会认为美国的生活是水深火热的（他们的最终观点 q ：美国就是每天到处发生枪击案），另一部分人最终会认为美国的空气都是甘甜的（他们的最终观点 $-q$ ：国内媒体不可信，只报道美国的坏事，不存在那么多枪击案），而很少会有人认为美国的生活没有那么好也没有那么坏（当然我不知道真实情况是哪种）。

PROPOSITION 1. *If $q_t = q$ for all $t \geq 0$, then user preference p_t converges to q if $p_0^\top q > 0$ and $-q$ if $p_0^\top q < 0$. Furthermore, the trajectory $(p_t)_{t \geq 0}$ lies in the cone defined by p_0 and $\pm q$. In particular, trajectory is characterized by:*

$$(p_t^\top q)^{-2} = 1 + \gamma_t^2 ((p_0^\top q)^{-2} - 1), \quad \gamma_t^2 = \begin{cases} (\eta + 1)^{-t} & \eta_t = \eta \\ \prod_{k=0}^{t-1} \frac{\eta^{-1}}{t+s+k} & \eta_t = \frac{\eta}{t+s} \end{cases} \quad (1)$$

结论2：可以做到个性化推荐时用户的偏好尽量不变

- 负面结论：前面的简单策略很可能无法保证用户的最终偏好和初始偏好一致。原因是无法在内容集合中找到和用户初始偏好一样的内容 ($q \neq p_0$)。定理描述如下 (其中 $r_t = p_0^T p_t$, $R(T) = \sum_{t=0}^{T-1} r_{\star} - r_t$)

PROPOSITION 5. Under fixed recommendations $q_t = q \neq p_0$, $R(T) \geq CT$ for a constant $C > 0$ depending on $p_0^T q$ in both the constant and decreasing step size settings.

- 正面结论：当内容集合比较丰富 ($p_0 \in \text{span}(\{q \mid q \in \mathcal{Q}, p_0^T q > 0\})$) 存在随机算法保证用户的最终偏好和初始偏好一致。

COROLLARY 7. Consider the setting and assumptions of Corollary 6 and further assume that support of the randomization is restricted to a set of items which is self-aligned with respect to v_1 and p_0 . Then, the regret of this recommendation approach is bounded with probability $1 - \delta$,

$$R(T) \leq 4e^2(s+2) \left(1 + \log\left(\frac{T+s}{s}\right)\right) + T\|v_1 - p_0\|$$

- 简要介绍一下这个随机算法：

$$\tilde{p}_{t+1} = p_t + \eta_t p_t^T q_t \cdot q_t \quad \xrightarrow{\text{两边取期望}} \quad \mathbb{E}[\tilde{p}_{t+1}] = (I + \eta_t \mathbb{E}[qq^T])p_t.$$

1. 获取 p_0 ,

2. 设计一种算法让 p_0 是 $\mathbb{E}[qq^T] = \sum_{i=1}^N \alpha_i q_i q_i^T$ 的最大特征值对应的特征向量, 其中 α_i 是每个时刻将 q_i 推荐给 p_0 的概率

$$\tilde{p}_{t+1} = p_t + \eta_t p_t^\top q_t \cdot q_t$$

结论3：若用户初始偏好未知，如何通过观测到的用户 和平台之间交互来确定用户偏好

- 当平台想要观测用户的初始偏好时，唯一的途径就是给用户推荐一个内容，通过用户的反应来学习用户的偏好。但是这种观测本身就会影响到用户的偏好。类似于量子力学中的观测者效应。
- 作者证明了存在算法可以解决这个问题，只要推荐的内容序列满足： $If q_0, \dots, q_{T-1} \text{ span } \mathbb{R}^d$
- 这个问题可以做的原因是我们假设用户偏好的演化严格服从这个公式： $\tilde{p}_{t+1} = p_t + \eta_t p_t^\top q_t \cdot q_t$

评价

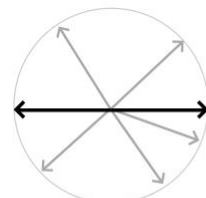
总结

- 建模推荐系统的时候避开了神经网络这个盲盒。
- 在建模了推荐系统对个体偏好的影响的基础上，探究了推荐系统和个体的交互过程。
- 从最大化用户满意程度、最大化用户兴趣不改变程度两个角度设计了算法。
- 发现固定推荐会导致人群观点极化。
- 设计算法用反馈信息还原用户初始偏好。

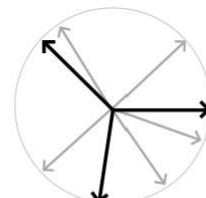
未来研究方向

- 建模考虑社交网络的基础上，以上结论和算法的性能保证还成立么？
- 实验验证。这篇论文纯理论，没有实验验证，当然做实验也比较困难。
- 将此理论结果和实际的推荐系统结合，设计出更棒的推荐机制。

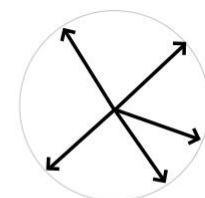
While non-personalized consumption leads to polarization [HJMR19, GKT21], personalized recommendations may lead initial preference (grey) to collapse to a subset of \mathcal{Q} . However, randomization can keep preferences stationary.



polarization



mode collapse



stationarity