



西安交通大学
XI'AN JIAOTONG UNIVERSITY

运筹学大作业报告

P3: 智能建筑群体热舒适感知

组长: 赵杰

组员: 黄伟鹏 张博雅

贾浚源 颜骁俊

目录

一、 问题背景与分析	2
1.1 问题背景	2
1.2 问题分析	2
二、 模型假设	3
三、 符号说明	4
四、 模型建立	4
五、 模型求解	8
5.1 确立模型参数	8
5.2 划分人群	11
六、 模型评价与改进	17
6.1 模型优点	17
6.2 模型缺点	17
6.3 模型改进	17
七、 参考文献	18
附录	19

智能建筑群体热舒适感知模型

摘要

近年来，信息物理融合技术的进步，推动着现代建筑朝着更智能化的方向发展。传统的建筑领域，往往更着重于对物理系统的考量，对人的舒适性缺乏关注。而现代智能建筑，将为居民提供更舒适的环境作为首要运行目标。热舒适是对热环境的一种主观满意度评价。当有多人处在同一办公或生活环境中时，由于其身体体质和环境偏好等不同，会产生不同的热需求。本文通过对已知数据的分析，建立热舒适度感知模型，明确人群类别划分的方法，并以 22 名受试人员为例，将其划分成四类，获取群体最大热舒适度。

首先，想要制定人群类别划分策略，就需要建立对热舒适度的二分类模型，来确定同一受试者在确定的环境温度、环境湿度以及自身热偏好情况下的热舒适情况。我们根据单一变量原则结合体感温度和相对湿度的经验公式，初步建立单个人的热舒适感知模型。然而，通过对数据的进一步分析，发现人的体感温度可能取到的范围可以被压缩至一个更小的区间，即对于偏好冷的人，我们只需要关注其舒适区间的上限，对于偏好热的人，我们只需要关注舒适区间的下限。针对同一个人的同一热偏好，我们发现舒适区间和不舒适区间有相当大的重合部分，于是引入**灰色地带**，确立最终的**热舒适度模型**。

其次，需要明确灰色地带的边界，使模型的正确率达到最高。我们以预测率为约束条件，正确率为目标函数，确定模型实现阶段的数学规划。由于直线的两个参数都没有确定，无法采用遍历的算法进行规划。因此，我们运用“两点连一线”的思想，将整个算法过程分为两步，先后确定灰色地带的两个边界，以达到最大的正确率。在程序实现过程中，我们发现大多数结果的正确率都在 80%、90% 左右，预测率也在可接受范围之内，可认为该算法的效果达到预期。

最后，需要根据群体最大热舒适度的要求实现**人群类别划分**。对于热偏好相同的人，我们将不同受试者的直线的近似程度进行比较，相似程度较大（生理指标对于热舒适度的影响较为接近）的归为一类。通过对结果的分析，我们从物理意义出发，以“包夹面积”最小为目标，借用“哈夫曼树”的算法思想，对不同热偏好的受试者进行初次划分，并进行节点的调整，最终使得 4 个房间的人数分布为 5, 5, 6, 6，实现了人群的动态划分。

关键词：热舒适度感知模型 人群类别划分 灰色地带

一、 问题背景与分析

1.1 问题背景

在全球工业化与城市化进程快速发展的大背景下,建筑能源需求总量急剧增加,在发达国家,建筑能耗占国家总能耗的 30%~40%。我国建筑能耗在全国能源消费总量的占比为 21.11%左右^[1,2],已与工业耗能、交通耗能并列,成为我国能源消耗的三大领域。特别是公共建筑,占建筑能耗的 38.37%^[1],我国建筑能耗构成比例如图 1-1 所示。事实上,暖通空调(HVAC)系统是公共建筑物内的主要能源消耗者,占全球能源使用量的 30%以上^[3]。在未来的十年里,建筑能耗伴随着建筑总量的不断攀升和居住舒适度的提升,在全社会总能耗的占比还将进一步增加,或将超越工业而成为用能的第一领域,我国建筑空调能耗的“黑洞”也将进一步扩大。因此,必须实行对建筑能耗高效的优化控制,提高建筑物的能源效率和利用率,这对我国整体的可持续发展具有重要意义^[4,5]。

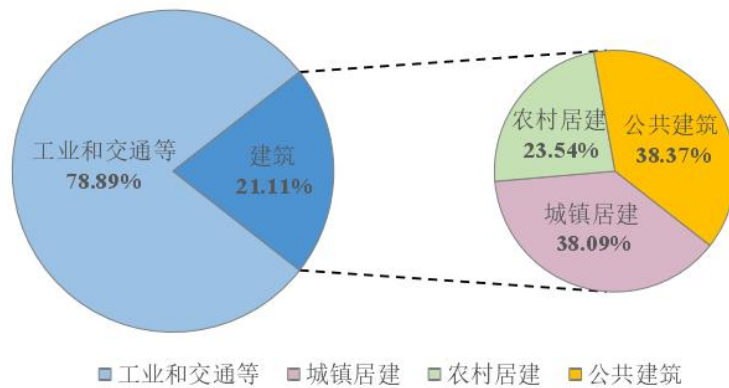


图 1-1 2017 年中国建筑能耗构成比例

1.2 问题分析

假设建筑室内环境参数由**室内温度**和**相对湿度**组成,在不同房间内有不同的人员办公,人员在环境内的热舒适性简化为**舒适与不舒适**的二分类问题。本题要求根据室内人员的**性别、年龄、身高、体重、热环境偏好**等参数,进行合理的人群类别划分,进而求得当前房间内群体热舒适度,并利用附件中的室内温度、相对湿度、受试者性别、年龄、身高、体重、热舒适等数据,以**最大群体热舒适度为目标,制定人群类别划分策略**。

首先,要制定人群类别划分策略,就应明确个体感到舒适的环境条件,也即需要建立对热舒适度的二分类模型,该模型能确定同一受试者在确定的环境温度、环境湿度以及自身热偏好情况下的热舒适情况。对于该模型,设定受试者的热舒适度为因变量,环境温度与环境湿度等环境参数,受试者的性别、年龄、身高、体重等生理指标,以及热偏好情况都是自变量。由于环境温度与湿度均为自变量,无法讨论二者之间的相关关系,利用回归分析等方法。同时,个体的热偏好也是自变量,讨论个体热偏好改变的原因过于复杂,本文不予研究。

要明确个体感到舒适的环境条件，理应严格控制变量，遵循单一变量原则，研究同一个人同一热偏好下对于不同的温度与湿度所感受到的热舒适度。同时，题目中的“等参数”也指出实际上影响热舒适感知的参数远远不止上文所提参数，为了是模型更加精简，我们只考虑上述主要参数的影响，来建立热舒适感知模型；最后，群体最大热舒适的目标也要求我们更加关注一个人感觉舒适的区域，而不是一个人感觉不舒适的区域。图 1-2 是本小组解决问题的步骤说明。

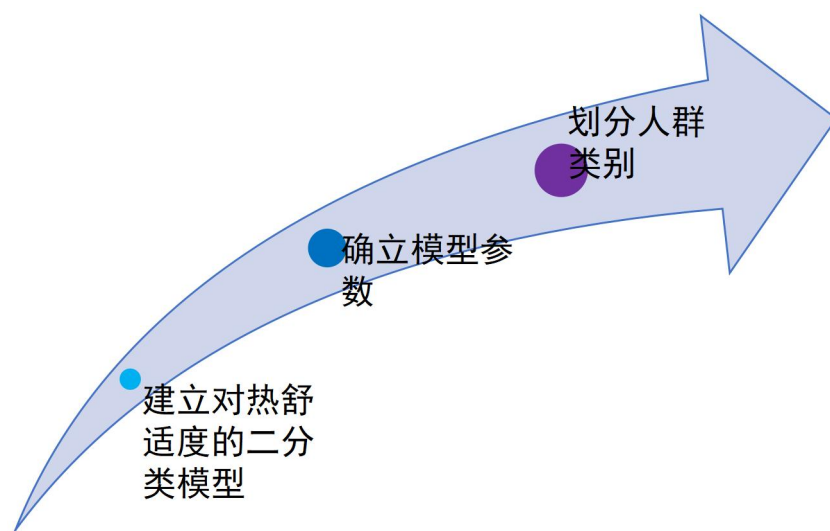


图 1-2 模型实现步骤

二、 模型假设

- 1、体感温度会受到环境温度和环境湿度的综合影响，在环境温度不变的情况下，湿度每增加 10%，体感温度下降某一具体值。
- 2、同一受试者在体感温度舒适区间的上下限附近存在难以预测热舒适度情况的灰色地带。
- 3、不同受试者的生理指标对于体感温度的影响不会导致函数关系的改变，函数参数仅由热偏好和个体生理指标影响。
- 4、在同一热偏好情况下，生理指标对于函数参数的影响是充分且均匀的，不会出现生理指标对参数的突变影响。
- 5、环境参数对于体感温度的影响是以环境温度与环境湿度项之间简单的加和关系来体现的。

三、 符号说明

符号	说明	单位
T_a	体感温度	℃
T	环境温度	℃
H	湿度	%
$accuracy$	正确率	/
$forecast$	预测率	/

四、 模型建立

由于环境因素主要是受**室内温度**和**相对湿度**影响,根据单一变量原则,我们控制性别、年龄、身高、体重、热环境偏好等参数不变,只改变相对湿度和环境温度,研究**同一个人**在**同一热偏好**下的热舒适区间(需要注意到在不同的生理条件状况下,同一个人对于相同的环境可能会体现出不同的热偏好。比如,对于同一个人,刚运动完后进入空调房和在办公室久坐一上午后进入空调房,就会呈现出不一样的热偏好)。据此,结合体感温度和相对湿度的经验公式,其中 $f(H)$ 为以相对湿度为自变量的函数

$$T_a = T + f(H) \quad (1)$$

而体感温度是指人体所感受到的冷暖程度,转换成同等之温度,会受到**气温**、**风速**与**相对湿度**的综合影响。以湿度为例,湿度每增加 10%,体感温度则降低 1℃。根据这一条件,我们猜测湿度每增加 10%,体感温度则降低某个具体数值 C,对于不同人, C 的值不同,由此可以体现一个人对湿度变化的耐受程度。可以推导出 $f(H)$ 的表达式

$$T_{a1} - T_{a2} = T + f(H) - T - f(0.9H) = f(H) - f(0.9H) = C \quad (2)$$

从 $f(H) - f(0.9H) = C$ 中,可以推测出 $f(H)$ 以对数函数为主要形式,因此假设 $f(H) = a * \ln H + b$, a 与 b 的值和每个人的特性有关,反映了一个人对于温度变化的耐受程度,可以以此为人群划分的依据。由此可得体感温度表达式(3),同时建立个人热舒适感知模型,如式(4)。

$$T_a = T + a * \ln H + b \quad (3)$$

$$G(T_a) = \begin{cases} 1(\text{绿色}), T_a \in \Omega_1 \\ -1(\text{红色}), T_a \notin \Omega_1 \end{cases}, \text{ 其中 } \Omega_1 \text{ 为热舒适区间。} \quad (4)$$

按照常理，一个具有特定偏好的人的舒适区间和不舒适区间应该如下图所示，图 4-1 从上至下分别为偏好冷的人、偏好热的人、偏好中性的人的舒适与不舒适区间。其中绿色代表舒适，红色代表不舒适。我们猜测热偏好应会大幅影响舒适区间的位置，如偏好冷的人的舒适区间将会在体感温度较低区域，偏好热应在体感温度较高区域，而偏好中性应在中间位置。而每个人的特性（即身高体重等各项参数）将会在小范围内影响舒适区间的位置。对于每个人来说，都应明确其舒适区间的上下限。

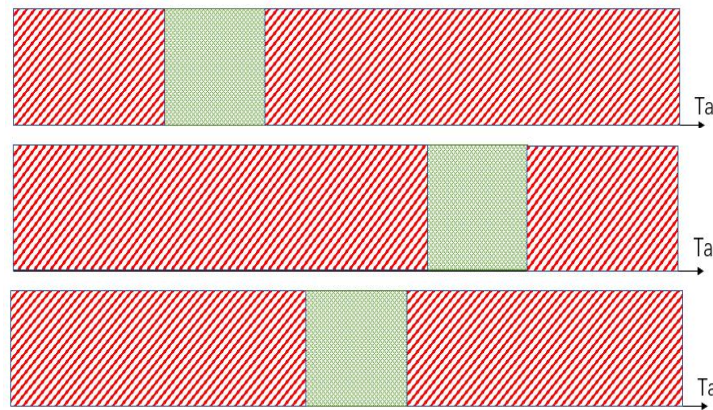


图 4-1 特定偏好人群的理论舒适与不舒适区间

对图 4-1 进行进一步分析，可以看出在极端体感温度（极大或极小）下，三种热偏好的人都应感到不舒适，因此若将所有感到不舒适数据集合起来，虽然体感温度受到相对湿度影响，但是在大量数据的情况下，温度分布直方图的上下极值对于三种热偏好来说应该相同，并且热偏好为中性的人理应有更多感到不舒适的数据。

然而，将所有人感到不舒适的数据按照热偏好区分，以温度为坐标轴作分布直方图，如图 4-2，二者在温度上极值差异较大，同时发现 1076 条偏好中性的数据中只有 2 条感到不舒适的，与上述的讨论不符，因此我们需要对模型进行修改和调整。

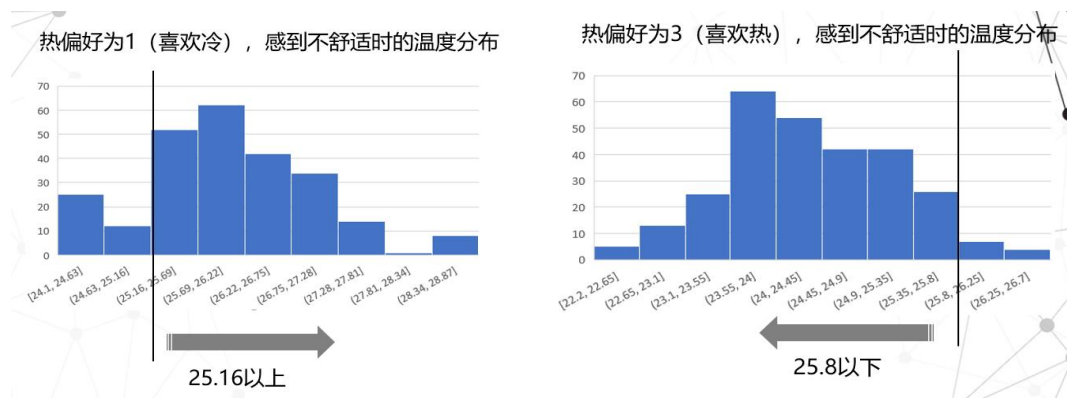


图 4-2 (a) 偏好冷的人群统计图

图 4-2 (b) 偏好热的人群统计图

实际上,根据题中数据给出的室温和相对湿度的大致分布,人的体感温度可能取到的范围可被压缩至一个更小的区间,计为区间 Φ ,即体感温度的可能取值区间。于是我们猜测每个人的舒适区间范围较大,如图4-3所示,为区间 Φ 与热偏好中性的人的舒适区间的关系。可以看出舒适区间完全包含了区间 Φ ,因此在正常情况下热偏好中性的人总是感到舒适,体现为1076条数据中只有两条数据感到不舒适。因此可将热偏好中性的人的热舒适感知模型由式(4)改为式(5)。由此我们也可认为只需要研究热偏好为冷和热的人。

$$G(T_a, \text{热偏好为中性}) = \begin{cases} 1(\text{绿色}), & \text{正常情况} \\ -1(\text{红色}), & \text{异常情况} \end{cases} \quad (5)$$

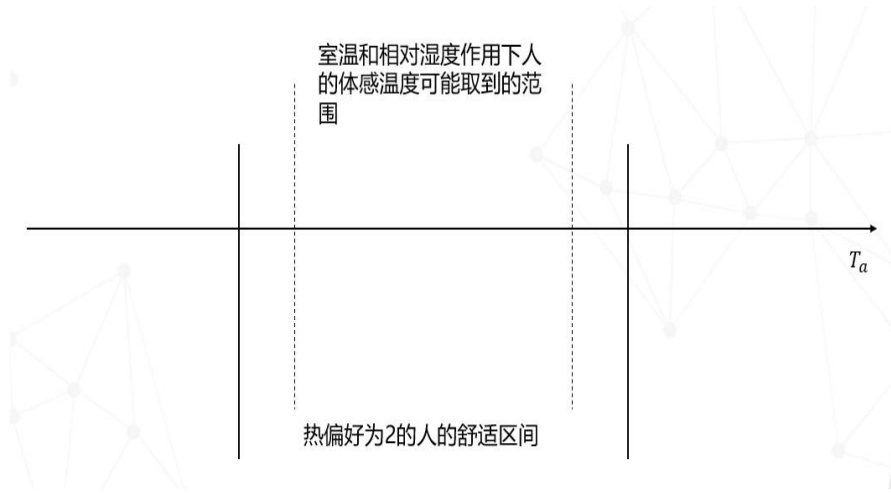


图4-3 区间 Φ 与热偏好中性的人的舒适区间

同样,如图4-4,可以看到由于热偏好冷的人由于舒适区间位于体感温度较低区域,且舒适区间较大,正常数值取不到舒适区间下限部分,当温度取极小的情况下,热偏好冷的人感到舒适,热偏好热的人感到不舒适,也就产生了上述讨论中的极值不同的情况。因此,对于热偏好冷的人,我们只需要关注其舒适区间的上限,对于热偏好热的人,我们只需要关注舒适区间的下限,由此可将个人热

舒适感知模型修改为 $G(T_a) = \begin{cases} 1(\text{绿色}), T_a > T_1 \\ -1(\text{红色}), T_a \leq T_1 \end{cases}$, 即

$$G(T_a, \text{热偏好3}) = \begin{cases} 1(\text{绿色}), T_a > T_1 \\ -1(\text{红色}), T_a \leq T_1 \end{cases} \quad (6)$$

$$G(T_a, \text{热偏好1}) = \begin{cases} 1(\text{绿色}), T_a < T_1 \\ -1(\text{红色}), T_a \geq T_1 \end{cases} \quad (7)$$

代入体感温度表达式(6), 通过化简可得

$$G(T_a, \text{热偏好3}) = \begin{cases} 1 \text{ (绿色)}, T_a > T_1 \\ -1 \text{ (红色)}, T_a \leq T_1 \end{cases}$$

$$\Rightarrow G(T, H, \text{热偏好3}) = \begin{cases} 1 \text{ (绿色)}, T + a * \ln H + b > T_1 \\ -1 \text{ (红色)}, T + a * \ln H + b \leq T_1 \end{cases}$$

$$\Rightarrow G(T, H, \text{热偏好3}) = \begin{cases} 1 \text{ (绿色)}, T > T_1 - a * \ln H - b \\ -1 \text{ (红色)}, T \leq T_1 - a * \ln H - b \end{cases}$$

$$\Rightarrow G(T, A, \text{热偏好3}) = \begin{cases} 1 \text{ (绿色)}, T > a' * A + b' \\ -1 \text{ (红色)}, T \leq a' * A + b' \end{cases} \quad \text{设} \ln H = A$$

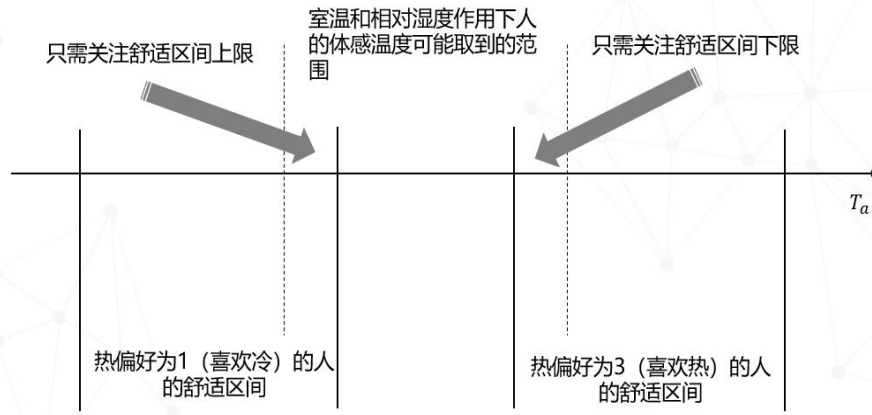


图 4-4 调整后的体感温度区间及特定偏好舒适区间

接下来针对同一个人的同一热偏好，以 T 为纵坐标，以 A 为横坐标作图，其中绿色代表舒适，红色代表不舒适，如图 4-5，来验证模型的准确性。图中两条直线是通过后面的算法处理出的结果。通过观察图像，以图 4-5 (a) 为例，在上面那条直线上方的点都是舒适点，在下面那条直线下方的点都是不舒适点，符合模型预测，但在两条直线中间部分，红点和绿点掺杂，不符合模型预测。

因此，需要再次对模型进行改进，引入灰色地带^[6]概念，即在灰色地带中的点的舒适度无法判断。在物理意义上解释为：在某些温度和湿度下，同样的人在刚运动完和没运动过的情况下，舒适与否是不同的，此时环境因素不是决定舒适与否的主要因素，因此无法通过环境参数预测舒适度，也就成为了灰色地带。

引入灰色地带后，热偏好为 1 和 3 的人的热舒适感知模型可以进一步修正为式 (8) 与式 (9)。

$$G(T_a, \text{热偏好3}) = \begin{cases} 1 \text{ (绿色)}, T > a' * A + b_1' \\ \text{无法判断 (红绿夹杂)}, a' * A + b_1' \leq T \leq a' * A + b_2' \\ -1 \text{ (红色)}, T < a' * A + b_2' \end{cases} \quad (8)$$

$$G(T_a, \text{热偏好} 1) = \begin{cases} 1(\text{绿色}), T < a' * A + b_1' \\ \text{无法判断(红绿夹杂)}, a' * A + b_1' \leq T \leq a' * A + b_2' \\ -1(\text{红色}), T > a' * A + b_2' \end{cases} \quad (9)$$

对比式(6)、式(7)与图 4-5 的结果，可以认为该模型的性能较好，上述猜测成立。

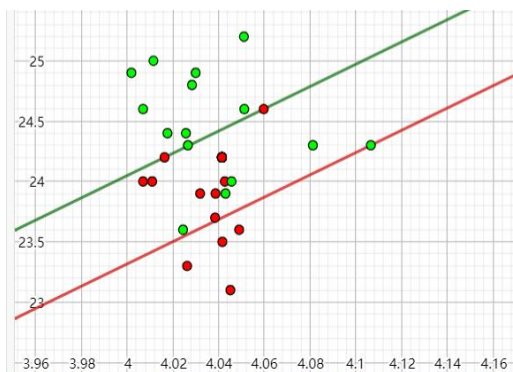


图 4-5 (a) 偏好冷的人的散点图

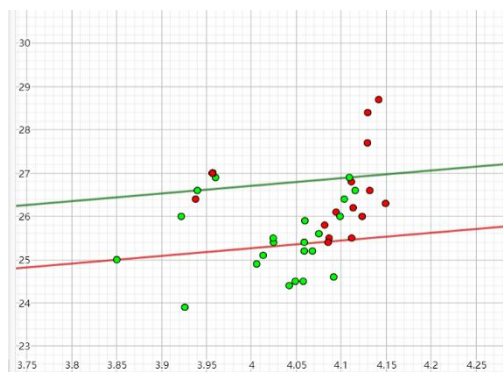


图 4-5 (b) 偏好热的人的散点图

五、 模型求解

5.1 确立模型参数

通过上述分析，确定个人热舒适感知数学模型为式(8)、式(9)。对于每个人来说，个人舒适感知模型中有三个未知参数 a 、 b_1 、 b_2 ，下面以热偏好为 3 时为例，确定这三个参数的具体过程。（热偏好为 1 时方法类似）

如图 5-1，通过确定三个未知参数 a 、 b_1 、 b_2 ，即确定两条直线，明确灰色地带的边界，使模型的正确率达到最高。

式(10)、式(11)分别给出正确率与预测率的定义表达式。

$$\text{正确率 } accuracy = \frac{b_1 + a_2}{a_1 + b_1 + a_2 + b_2} \quad (10)$$

$$\text{预测率 } forecast = \frac{a_1 + b_1 + a_2 + b_2}{n} \quad (11)$$

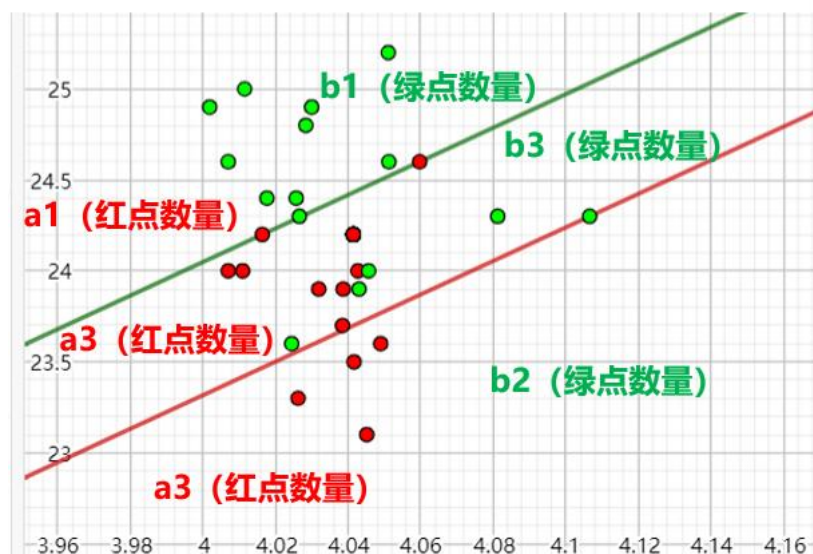


图 5-1 偏好热的人的灰色地带

预测率代表了除灰色地带以外的点的数量占点的总数的比例；正确率代表了除灰色地带以外的点中，物理意义正确的点占总点数的比例。由于在个人热舒适感知模型中，灰色地带中的点的舒适度的值无法预测，因此我们总希望模型的预测率应较高，体现在预测率应高于某一个预定值，因此预测率成为模型实现中的一个约束条件。而正确率则反应了模型的性能，我们总希望模型的正确率能达到最高，因此正确率成为模型实现中的一个目标值。但是该正确率由于忽略了灰色地带中的点，因此会产生相应问题。

由预测率以及正确率，可以确定模型实现阶段的一个数学规划，如式(12)。式中参数 a 为我们对于预测率的一个最低的预定值，在具体实现的过程中需要调节 a 的参数值，使模型达到更好的结果。

$$\begin{cases} \max z = accuracy \\ s.t. \text{ forecast} \geq a \end{cases} \quad (12)$$

由于图 5-1 中横纵坐标代表的变量都是自变量，二者之间没有相关关系，因而不能直接对图中的点采取回归分析、近似拟合等方法，而需要采用特殊方法。

通过大量作图与数据分析，发现灰色地带中的点的温度与相对湿度大多处于适中的位置，而如图 4-2，温度与相对湿度大多在适中位置，即灰色地带中的点的数量通常较多，意味着很可能有两个点落在灰色地带边界上。因此，连接灰色地带中的两个点，以此作为灰色地带的边界，在过灰色地带中的另一个点作该边界的平行线，以此作为灰色地带的另一个边界，这一组边界确定后，相应的正确率、预测率也随之确定，遍历灰色地带中的所有点，再从中选取使数学规划(12)达到最优的一组边界线，即认定其为所需的边界线。

该算法具有缺陷，首先需要保证在灰色地带中的点的数量够多，这便要求模型的预测率不能太高；同时，我们无法事先确定哪些点在灰色地带中，也就无法遍历灰色地带中的点，于是只能遍历所有的点，但是遍历所有的点所得的边界线组中总会存在灰色地带太大，正确率极高而预测率略大于 a 的情况（也就是预测率较小），这就对参数 a 的要求很高，很难实现。

因此对算法进行适当改进，将整个过程分成两步进行，如图 5-2 (a)，第一步先在所有点中挑选出两个点并连接，计算出相应的正确率，遍历所有的两点组合，挑选出直线上方绿点与下方红点总数最多的一组，以其连线作为灰色地带的一个预测边界。算法的第一步在计算正确率时考虑了所有点，因此保证了预测率，使其达到一个较高水平，也使整个算法对于 a 的要求降低。

第二步，如图 5-2 (b)，在所有点中挑选一个点，过该点作第一步所得边界的平行线，并遍历得出满足数学规划式 (9) 的约束条件，并使正确率达到最高的点，以该直线作为灰色地带另一边边界。通过两个步骤完成数学模型的实现。

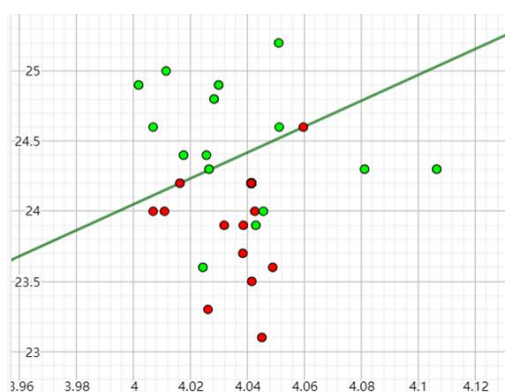


图 5-2 (a) 第一步算法

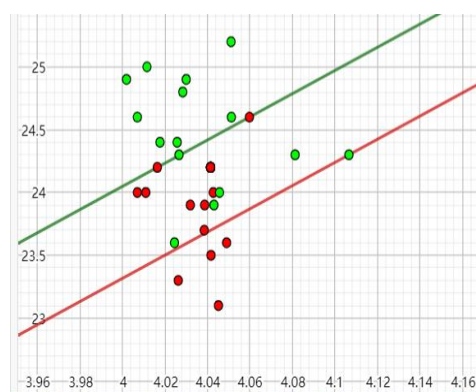


图 5-2 (b) 第二步算法

相应的 MATLAB 代码见附录 1，在程序运行的过程中，调节参数 a ，发现当参数 a 取 0.25 时效果较好，图 5-3 是部分结果图。

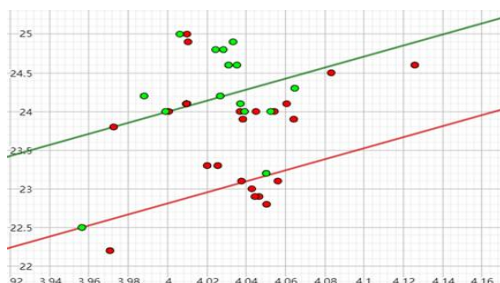
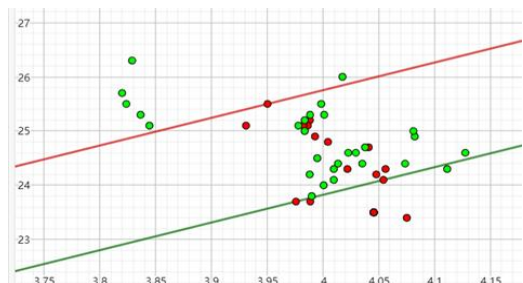
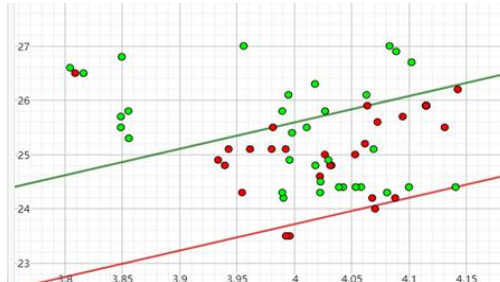
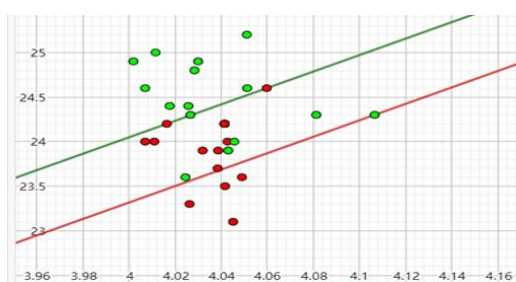


图 5-3 a 取 0.25 时结果图

如表 5-1 所示，大多数结果的正确率都在 80%、90%左右，预测率不高也不低，在可接受范围之内，可认为该算法的效果较好，同时所得图像也可反映出第二部分所建立的个人热舒适感知模型效果较好。因此可由此算法所得的参数 k、b1、b2 进行后续的人群类别划分。其中有部分数据太少，可认为此人在正常情况下不会出现该种热偏好。

新编号	人员编号	热偏好3								热偏好1								热偏好2	
		k	b1	b2	正确率	预测率	不舒适 (灰色地带)	舒适 (灰色地带)	数量	k	b1	b2	正确率	预测率	不舒适 (灰色地带)	舒适 (灰色地带)	数量		
1	1000000	2.9822	12.1424	13.5588	0.725	0.5634	0.4516	0.5484	71	4.7767	7.9812	5.0467	0.8333	0.3158	0.4615	0.5385	19	70	
2	3118105194	8.7181	-11.576	-12.082	0.9565	0.7931	0.1667	0.8333	29	1.2786	20.3949	20.9531	1	0.7143	0.125	0.875	28	15	
3	3118105205	6.6163	-3.3023	-4.1728	1	0.9024	0.25	0.75	41	2.7221	14.5194	14.5654	0.9412	0.8095	0	1	21	5	
4	3118105213	8.989	-12.676	-13.001	0.8333	0.72	0.4286	0.5714	25	3.8122	10.0323	11.4726	0.75	0.3333	0.5	0.5	24	58	
5	3118105214	8.2209	-9.6477	-9.9246	0.9286	0.7	0.1667	0.8333	20	5.5203	4.3713	6.3996	0.5238	0.8077	0.2	0.8	26	62	
6	3118105318	4.8665	6.1254	4.2512	0.9524	0.3443	0.575	0.425	61	4.0826	10.8005	9.2212	0.5714	0.4375	0.5556	0.4444	16	47	
7	3118105319	5.4019	3.9868	5.9139	0.85	0.6452	0.8182	0.1818	61	2.2946	17.9629	16.5788	1	0.3333	0.75	0.25	6	78	
8	3118305213	数据太少且极端								数据太少且极端								1	
9	3118305456	数据太少且极端								9.1438	-10.998	-13.122	0.8235	0.3696	0.5517	0.4483	46	82	
10	3118305457	4.4416	7.744	7.9858	0.7308	0.7879	0.5714	0.4286	33	5.0589	4.728	3.5766	0.7547	0.7162	0.7619	0.2381	74	86	
11	3118305488	5.1033	4.6233	5.9257	0.7714	0.7143	0.5714	0.4286	49	7.6626	-6.0939	-7.2696	0.7586	0.7436	0.7	0.3	39	55	
12	3118305504	数据太少								数据太少								83	
13	3119105211	数据太少								2.9359	13.9112	13.4761	0.9091	0.6875	0.6	0.4	16	65	
14	4115005024	5.0976	3.4332	5.0611	0.8333	0.2727	0.375	0.625	44	数据太少								37	
15	4117005108	数据太少								数据太少								45	
16	4118005124	数据太少								6.896	-2.1734	-3.477	0.6571	0.5833	0.4	0.6	60	44	
17	4118014038	9.2229	-12.843	-13.575	0.9333	0.5357	0.6154	0.3846	28	数据太少								46	
18	4119105033	6.0863	0.1862	3.5605	1	0.4118	0.3	0.7	17	数据太少								21	
19	4119105049	数据太少或大极端								数据太少或大极端								58	
20	3117315007	7.1353	-4.5453	-5.7302	0.8333	0.4865	0.6842	0.3158	37	1.7727	19.616	18.1578	0.9286	0.3784	0.4783	0.5217	37	66	
21	4118005125	2.5598	14.5531	15.1804	0.75	0.4138	0.4706	0.5294	29	无数据								31	
22	4116005044	无数据								5.2334	4.0495	5.9148	0.8235	0.5	0.5294	0.4706	34	22	

表 5-1 算法结果呈现

5.2 划分人群

通过以上对模型的建立与实现，我们实现了对同一受试者在某一热偏好下，由其所处的环境湿度和环境温度预测其热舒适度，现要根据群体最大热舒适度的要求实现对人群类别的划分。

$$G(T, A, \text{热偏好3}) = \begin{cases} 1(\text{绿色}), T > a' * A + b_1' \\ \text{无法判断(红绿夹杂)}, a' * A + b_1' \leq T \leq a' * A + b_2' \\ -1(\text{红色}), T < a' * A + b_2' \end{cases} \quad (13)$$

$$G(T, A, \text{热偏好3}) = \begin{cases} 1(\text{绿色}), T < a' * A + b_1' \\ \text{无法判断(红绿夹杂)}, a' * A + b_1' \leq T \leq a' * A + b_2' \\ -1(\text{红色}), T > a' * A + b_2' \end{cases} \quad (14)$$

回顾热偏好为 1 和 3 的人的热舒适感知模型，即式(13)与式(14)，同时参考图 5-4 可以看出二者的绝对舒适地带，即可以预测且几乎感到舒适的地带分别为下半平面与上半平面，二者重叠部分很少，可以认为二者差异很大，不能将二者

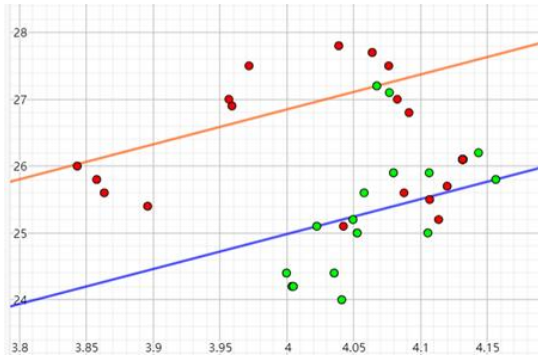


图 5-4 (a) 偏好冷的人的舒适区间分布
归为一类。

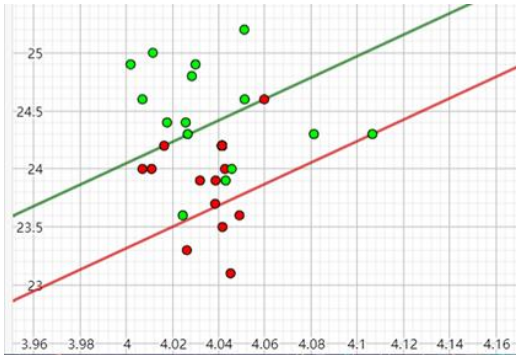


图 5-4 (b) 偏好热的人的舒适区间分布

因此，对于不同的热偏好情况，热舒适度的舒适区间有较大差别。为了满足群体热舒适度最大的目标，我们有必要在不同的热偏好情况下对人群进行分类，将热偏好偏冷和偏热的人群分开，再在同一热偏好下对人群进行细分。下面我们以热偏好为 3 的人群为例，分析如何在同一热偏好内进一步细分人群。

由于我们的目标是追求最大舒适度，而在灰色地带中的点无法保证其一定舒适，因此，对于热偏好冷（如图 5-4 (a)），我们更加看重位于下方的直线，该直线下方是能保证舒适的区域。同样，对于热偏好热（如图），我们更加看重位于上方的直线。又由于灰色地带边界线的两个主要参数 a' 、 $b1'$ 是由受试者自身的生理指标决定的，所以可以对同一热偏好下，上述讨论中的直线便能作为人群划分的依据。将不同受试者的直线的近似程度进行比较，相似程度较大的归为一类，说明二者的生理指标对于热舒适度的影响较为接近。

正如模型建立部分所讨论的，由于室内空调的存在，温度与湿度是有一定区间限制的，不会出现较为离谱的数据。图 5-5 分别是温度与湿度的分布直方图，在这之中可以取出温度与湿度可能取到的极值，以这些极值做出矩形框，可以认为矩形框外的温度与湿度是取不到的，因此我们对于直线“相似程度”的讨论也应在这个矩形框内进行。

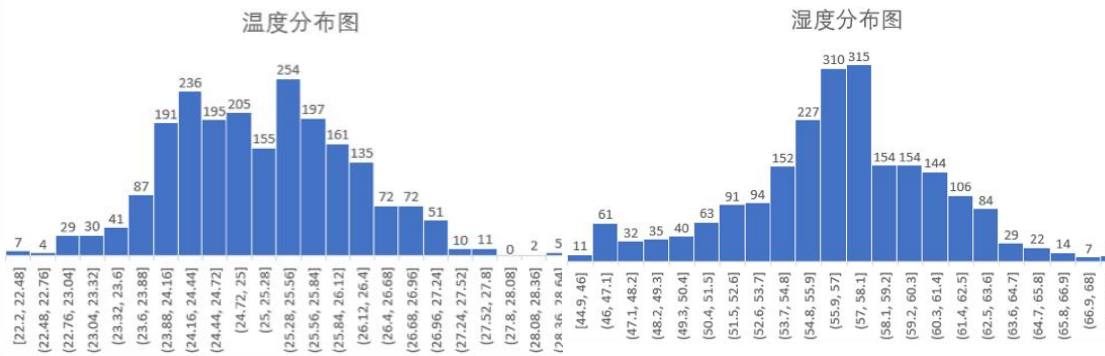


图 5-5 (a) 温度分布直方图

图 5-5 (b) 湿度分布直方图

理所当然，我们首先会想到通过比较两个参数的近似程度来确定两条直线的相似程度。然而，如图 5-6 所示的三条直线，6.45 与 6.1 的差距显然大于与 6 的差距，但是由于 -1.8 的引入，导致前二者的相似程度反而较大，说明直接由两个参数的近似程度来确定直线的相似程度是不可靠的。

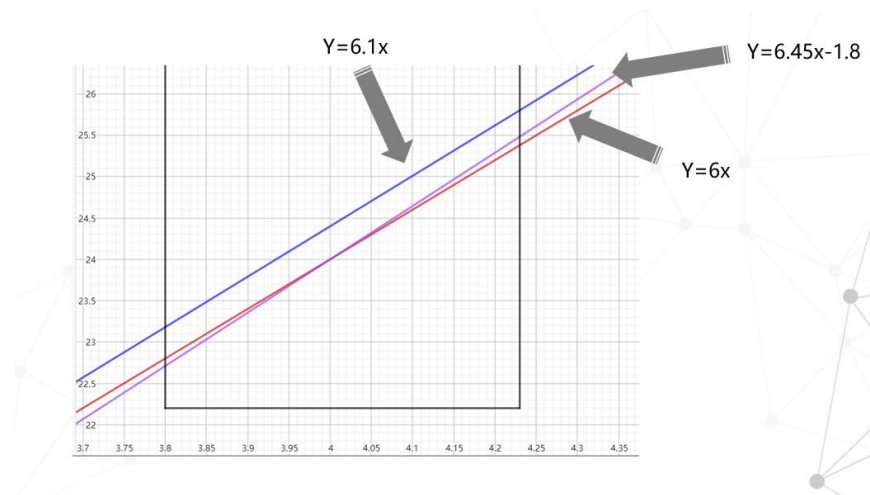


图 5-6 通过参数确定两条直线的相似程度

接下来从物理意义角度来确定两条直线的相似程度。仍以热偏好 3 为例，如图 5-7，对于这两个人来说，在两条直线上方以及下方的绿色区域所代表物理意义相同，即上方二者均感觉舒适，在下方二者共同处于灰色地带或者感觉不舒服。而图中的红色区域所代表的物理意义则不同，即一人感觉舒适，一人感觉不舒服。当两个人的特性足够相似时，红色部分的面积理应足够小。因此，认为红色区域最小时，两个人的相似程度最大，可归为一类。

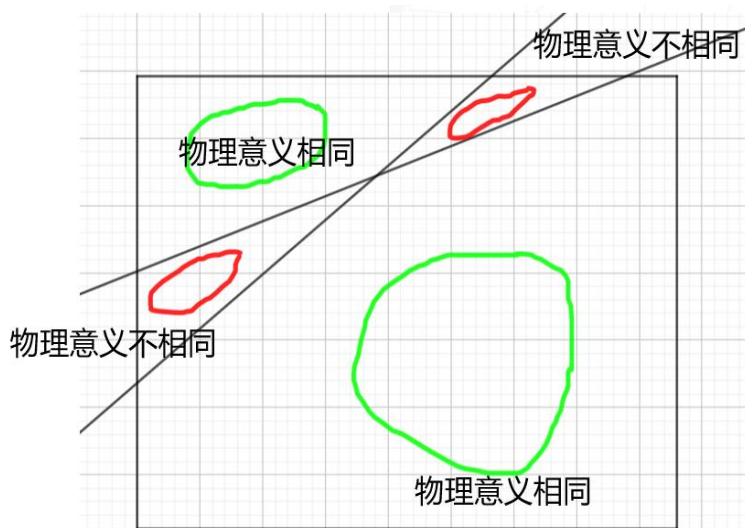


图 5-7 通过物理意义确定两条直线的相似程度

然而，在实际的算法实现中，由于两条直线交点的位置有 4 种可能，每条直线与矩形边界的交点在哪两条边的可能性 4 种，总的情况太多，给算法的实现带

来了很大的困难。因此，如图 5-8，采用红色四边形的面积近似替代红色区域的面积，并认为红色四边形的面积越小，两个人的相似程度就越高。

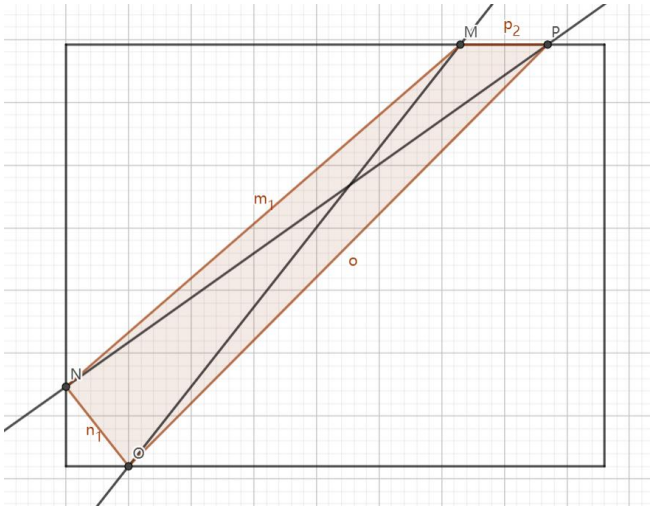


图 5-8 “近似”思想的应用

明确了两个人相似程度的定义，接下来便能进行人群类别划分。算法思路如下：如图 5-9，先在所有人群中挑选出相似程度最高的两个人，让它们组成一个新的节点，同时由于二者较为相似，以这两个人两条直线的角平分线作为新节点所携带的信息，完成一次挑选。再将新节点放回人群中，再次挑选出两个相似程度最高的人，重复操作。每一趟挑选都会减少一个节点，当节点只剩一个时算法结束，因此该算法肯定能结束。

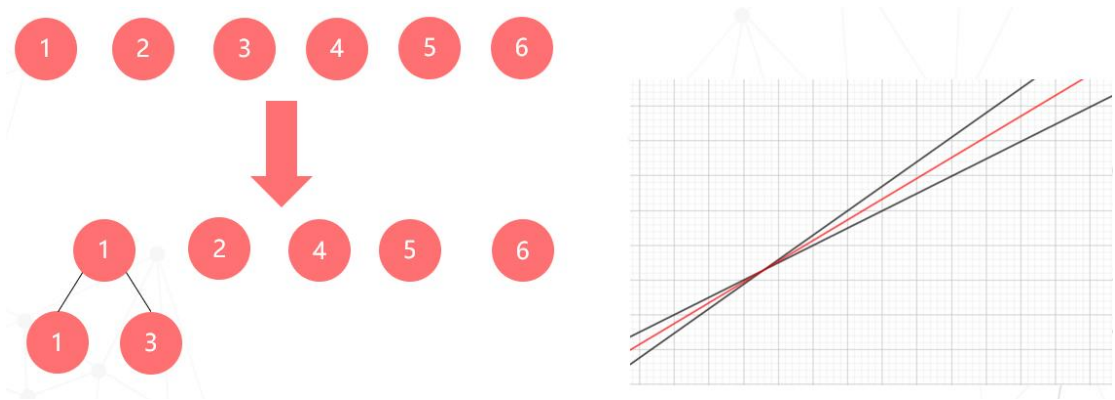


图 5-9 人群划分示意图

具体 MATLAB 代码见附录 2。通过代码所得人群的树状分类图如 5-10(a)(热偏好 3)、图 5-10(b)(热偏好 1) 所示。图中每个叶子节点代表一个人，从树

状图的根节点往下走，越靠近下方，两个人的相似程度越高，将其划分入同一人群的效果更好。

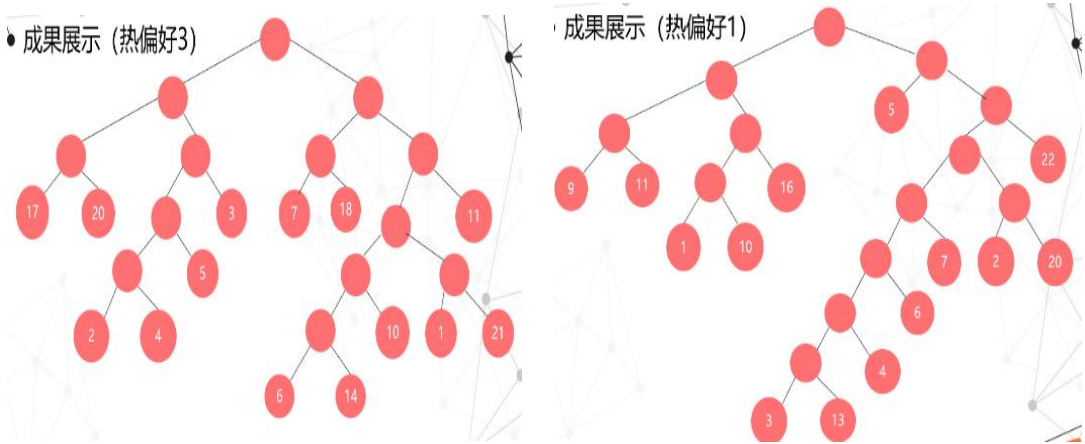


图 5-10 (a) 偏好热人群的树状分类图 图 5-10 (b) 偏好冷人群的树状分类图

为了验证该算法的效果，以热偏好 3 为例，作出分类后各组直线的具体图像，如下图 5-11 所示，可以看出各组直线在矩形框内的相似程度很高，因此可以认为该算法的效果较好。

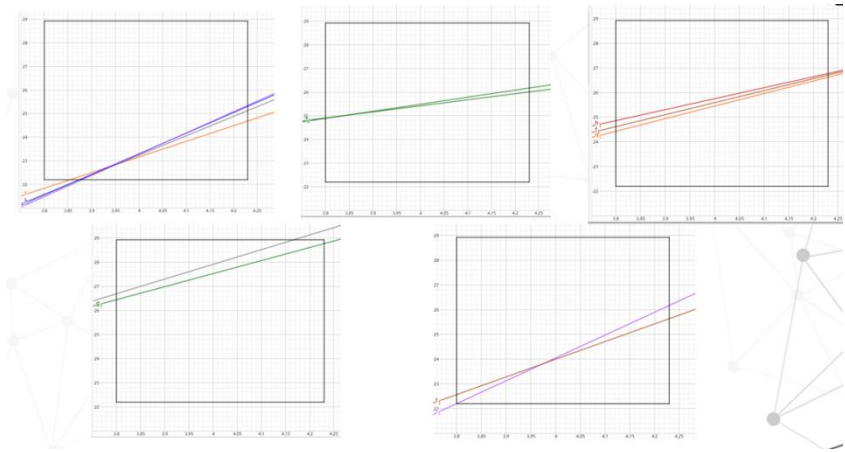


图 5-11 分类后各组图像

当然，从图 5-10 中不难发现有许多人的点重复了，这是因为一个人可以有多种热偏好，且热偏好也是自变量，其变化无法预测。但是实际上，在同一时刻，一个人只会表现出一种热偏好，因此我们不仅需要上述的静态人群划分，还需要借助上述静态人群划分完成进一步的动态人群划分，以实现最大群体舒适度的目标。以下通过具体例子讲述动态人群划分的思路。

当某一天员工到达公司，需要确定每位员工的热偏好，统计得出热偏好为 1 的员工为 2、4、5、9、10、11、13、16、20、22；热偏好为 2 的员工为 8、12、15、19；热偏好为 3 的员工为 1、3、6、7、14、17、18、21，现在进行静态树

的更改。将图 5-10 中不存在的节点剪除，得到两棵新的动态人群划分树，如图 5-12 所示。由于公司中一共有 4 个房间，应尽量使房间中的人数为 5、5、6、6

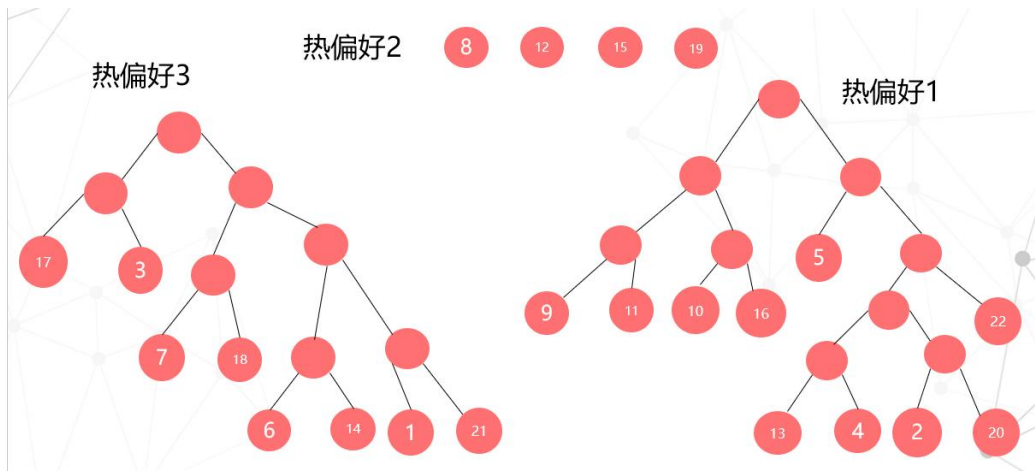


图 5-12 剪除划分树中不存在的节点

从两棵树的根节点出发，往下走一个节点，由于热偏好 2 的人在正常情况下都觉得舒适，因此无论将其放入哪个人群，都能提高人群的群体舒适度，类似于“万金油”的存在，由于两棵树的左子树上的叶子节点的个数较少，因此用热偏好 2 的节点将人数补全，如图 5-13 所示，每个圆圈即代表一个房间，可看到房间中的人数刚好为 5, 5, 6, 6，即可认为实现了人群的动态划分。进行动态划分时，根据房间的总数，从根节点出发，以此往下走，若房间数较多，则往下走得更多，反之则走得少，再利用热偏好 2 的节点对人数进行补全，便能借助静态人群划分树状图实现动态人群划分。

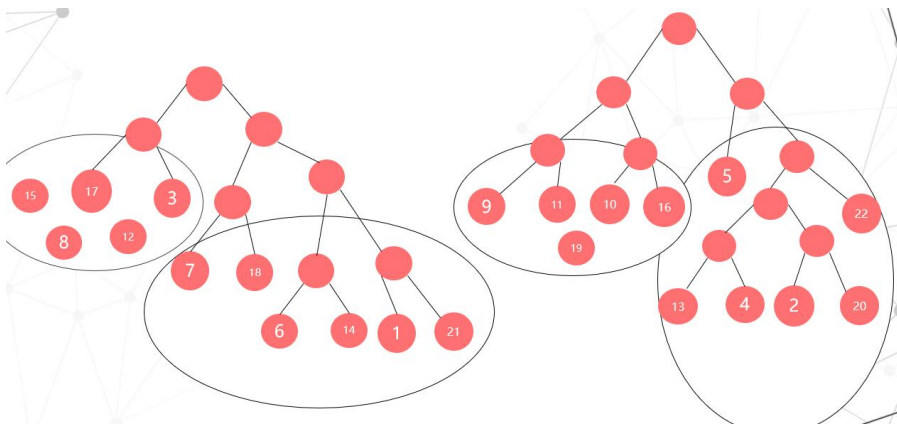


图 5-13 新的动态人群划分树状图

六、 模型评价与改进

6.1 模型优点

- 1、建立热舒适度模型时，通过对海量数据的分析，将人的体感温度可能取到的范围可被压缩至一个更小的区间，使原本复杂的问题简单化。
- 2、由于同一个人的同一热偏好的舒适区间和不舒适区间有较大的重合，引入灰色地带，使模型符合预测。
- 3、确定模型参数时的算法第一步在计算正确率时考虑了所有点，因此保证了预测率，使其达到一个较高水平，也使整个算法对于 α 的要求降低。
- 4、为更准确地划分人群，将生理指标对于热舒适度的影响相似程度转化为可数据化的不同受试者的直线的近似程度。同时从物理意义角度而不是单纯从斜率和截距来确定两条直线的相似程度。
- 5、具体划分过程中，借鉴了数据结构中的“哈夫曼树”的算法思想，实现动态人群分类。

6.2 模型缺点

- 1、确立灰色地带的过程中，由于直线是通过任意两个点连接得到的，没有办法遍历所有可能情况的直线，因此得到的结果只能保证相对最优。
- 2、划分人群时，计算对应面积时采用了近似的思想，不能保证精确度，可能会对最终的划分结果产生一定的影响。
- 3、

6.3 模型改进

本模型采用的都是基于一些简单易实现的方法，在人群划分方面或许可以借鉴机器学习和深度学习中的某些算法，使得划分的结果更准确。

七、 参考文献

- [1] 中国建筑节能协会. 2019 中国建筑能耗研究报告[J]. 建筑, 2020, (07): 30-39.
- [2] 中国建筑节能协会. 2018 中国建筑能耗研究报告[J]. 建筑, 2019, (02): 26-31.
- [3] 赵丽丽. 我国建筑能耗现状分析[J]. 住宅与房地产, 2016, (30): 25.
- [4] Zhao D, McCoy A P, Du J, et al. Interaction effects of building technology and resident behavior on energy consumption in residential buildings[J]. Energy and Buildings, 2017, 134(JAN.): 223-233.
- [5] Hannan M A, Faisal M, Ker P J, et al. A Review of Internet of Energy Based Building Energy Management Systems: Issues and Recommendations[J]. IEEE Access, 2018, 6: 38997-39014.
- [7] 王立栋. 蛰伏于灰色地带之软件[J]. 新电脑, 2007(12):186-189.

附录

附录 1

介绍：热偏好模型实现代码

```
//热偏好 3 模型实现代码
n=37; %元素个数
a1=0;b1=0;a2=0;b2=0; %1 上 2 下, a 红 b 绿
accuracy=0;
for i=1:1:n-1
    for j=i+1:1:n
        a10=0;b10=0;a20=0;b20=0;
        if S1(i,2)==S1(j,2)
            continue;
        end
        k0=(S1(i,1)-S1(j,1))/(S1(i,2)-S1(j,2));
        b0=S1(i,1)-k0*S1(i,2);
        if k0<=0 || k0>10
            continue;
        end
        for i=1:1:n
            c=k0*S1(i,2)+b0;
            if S1(i,1)>c %点在直线上方
                if S1(i,3)==-1
                    a10=a10+1;
                else
                    b10=b10+1;
                end
            else
                if S1(i,3)==-1
                    a20=a20+1;
                else
                    b20=b20+1;
                end
            end
        end
        accuracy0=(b10+a20)/n;
        forecast0=(a10+b10)/n;
        if accuracy0>accuracy && forecast0>=0.25
            accuracy=accuracy0;
            forecast=forecast0;
            k=k0;
            b=b0;
        end
    end
end
```

```

        if accuracy0==accuracy && forecast0>forecast
            accuracy=accuracy0;
            forecast=forecast0;
            k=k0;
            b=b0;
        end
    end
end
k
b
accuracy
forecast
accuracy=0;
forecast=0;
for i=1:1:n
    a10=0;b10=0;a20=0;b20=0;a30=0;b30=0;
    b2=S1(i,1)-k*S1(i,2);
    if b2==b
        continue;
    end
    for j=1:1:n
        c1=k*S1(j,2)+b;
        c2=k*S1(j,2)+b2;
        if S1(j,1)>c1 && S1(j,1)>c2 %点在直线上方
            if S1(j,3)==-1
                a10=a10+1;
            else
                b10=b10+1;
            end
        elseif S1(j,1)<c1 && S1(j,1)<c2
            if S1(j,3)==-1
                a20=a20+1;
            else
                b20=b20+1;
            end
        else
            if S1(j,3)==-1
                a30=a30+1;
            else
                b30=b30+1;
            end
        end
    end
end
accuracy0=(b10+a20)/(a10+b10+a20+b20);

```

```

forecast0=(a10+b10+a20+b20)/n;
if accuracy0>accuracy && forecast0>=0.25
    accuracy=accuracy0;
    forecast=forecast0;
    b22=b2;
    a3=a30;
    b3=b30;
end
if accuracy0==accuracy && forecast0>forecast
    accuracy=accuracy0;
    forecast=forecast0;
    b22=b2;
    a3=a30;
    b3=b30;
end
end
b22
accuracy
forecast
a3/(a3+b3)
b3/(a3+b3)

//热偏好 1 模型实现代码
n=28; %元素个数
a1=0;b1=0;a2=0;b2=0; %1 上 2 下, a 红 b 绿
accuracy=0;
for i=1:1:n-1
    for j=i+1:1:n
        a10=0;b10=0;a20=0;b20=0;
        if S1(i,2)==S1(j,2)
            continue;
        end
        k0=(S1(i,1)-S1(j,1))/(S1(i,2)-S1(j,2));
        b0=S1(i,1)-k0*S1(i,2);
        if k0<=0 || k0>10
            continue;
        end
        for i=1:1:n
            c=k0*S1(i,2)+b0;
            if S1(i,1)>c %点在直线上方
                if S1(i,3)==-1
                    a10=a10+1;
                else
                    b10=b10+1;
                end
            end
        end
    end
end

```

```

        end
    else
        if S1(i,3)==-1
            a20=a20+1;
        else
            b20=b20+1;
        end
    end
end
accuracy0=(b20+a10)/n;
forecast0=(a20+b20)/n;
if accuracy0>accuracy && forecast0>=0.25
    accuracy=accuracy0;
    forecast=forecast0;
    k=k0;
    b=b0;
end
if accuracy0==accuracy && forecast0>forecast
    accuracy=accuracy0;
    forecast=forecast0;
    k=k0;
    b=b0;
end
end
end
k
b
accuracy
forecast
accuracy=0;
forecast=0;
for i=1:1:n
    a10=0;b10=0;a20=0;b20=0;a30=0;b30=0;
    b2=S1(i,1)-k*S1(i,2);
    if b2==b
        continue;
    end
    for j=1:1:n
        c1=k*S1(j,2)+b;
        c2=k*S1(j,2)+b2;
        if S1(j,1)>c1 && S1(j,1)>c2 %点在直线上方
            if S1(j,3)==-1
                a10=a10+1;
            else

```



```

        b10=b10+1;
    end
elseif S1(j,1)<c1 && S1(j,1)<c2
    if S1(j,3)==-1
        a20=a20+1;
    else
        b20=b20+1;
    end
else
    if S1(j,3)==-1
        a30=a30+1;
    else
        b30=b30+1;
    end
end
end
accuracy0=(b20+a10)/(a10+b10+a20+b20);
forecast0=(a10+b10+a20+b20)/n;
if accuracy0>accuracy && forecast0>=0.25
    accuracy=accuracy0;
    forecast=forecast0;
    b22=b2;
    a3=a30;
    b3=b30;
end
if accuracy0==accuracy && forecast0>forecast
    accuracy=accuracy0;
    forecast=forecast0;
    b22=b2;
    a3=a30;
    b3=b30;
end
end
b22
accuracy
forecast
a3/(a3+b3)
b3/(a3+b3)

```

附录 2

介绍：人群划分代码

```
n=14;
s=10000;
p=2;q=1;
y1=23.88;y2=26.4;y3=22.2;y4=28.92;
x1=3.94;x2=4.15;x3=3.8;x4=4.23;
for k=1:1:n-1
    s=10000;
    for i=1:1:n-1
        for j=i+1:1:n
            a1=0;a2=0;a3=0;a4=0;a5=0;a6=0;a7=0;a8=0;
            b1=0;b2=0;b3=0;b4=0;b5=0;b6=0;b7=0;b8=0;
            if S1(i,2)~=0 && S1(j,2)~=0
                if (y2-S1(i,3))/S1(i,2)>x1 && (y2-S1(i,3))/S1(i,2)<=x2
                    a1=(y2-S1(i,3))/S1(i,2);
                    b1=y2;
                elseif (S1(i,2)*x2+S1(i,3))>y1 && (S1(i,2)*x2+S1(i,3))<y2
                    a1=x2;
                    b1=S1(i,2)*x2+S1(i,3);
                end
                if (y1-S1(i,3))/S1(i,2)>=x1 && (y1-S1(i,3))/S1(i,2)<x2
                    a2=(y1-S1(i,3))/S1(i,2);
                    b2=y1;
                elseif (S1(i,2)*x1+S1(i,3))>y1 && (S1(i,2)*x1+S1(i,3))<y2
                    a2=x1;
                    b2=S1(i,2)*x1+S1(i,3);
                end
                if (y2-S1(j,3))/S1(j,2)>x1 && (y2-S1(j,3))/S1(j,2)<=x2
                    a3=(y2-S1(j,3))/S1(j,2);
                    b3=y2;
                elseif (S1(j,2)*x2+S1(j,3))>y1 && (S1(j,2)*x2+S1(j,3))<y2
                    a3=x2;
                    b3=S1(j,2)*x2+S1(j,3);
                end
                if (y1-S1(j,3))/S1(j,2)>=x1 && (y1-S1(j,3))/S1(j,2)<x2
                    a4=(y1-S1(j,3))/S1(j,2);
                    b4=y1;
                elseif (S1(j,2)*x1+S1(j,3))>y1 && (S1(j,2)*x1+S1(j,3))<y2
                    a4=x1;
                    b4=S1(j,2)*x1+S1(j,3);
                end
            end
        end
    end
end
```

```

if a1>a3 || b1<b3
    t=a3;a3=a1;a1=t;
    t=b3;b3=b1;b1=t;
end
if a2>a4 || b2<b4
    t=a4;a4=a2;a2=t;
    t=b4;b4=b2;b2=t;
end
k1=(b1-b2)/(a1-a2);k2=(b3-b2)/(a3-a2);k3=(b3-b4)/(a3-a4);
s1=(k1-k2)*((a1-a2)*(a3-a2)+(b1-b2)*(b3-b2))/(2*(1+k1*k2))+
(k3-k2)*((a2-a3)*(a4-a3)+(b2-b3)*(b4-b3))/(2*(1+k2*k3));
if (y4-S1(i,3))/S1(i,2)>x3 && (y4-S1(i,3))/S1(i,2)<=x4
    a5=(y4-S1(i,3))/S1(i,2);
    b5=y4;
elseif (S1(i,2)*x4+S1(i,3))>y3 && (S1(i,2)*x4+S1(i,3))<y4
    a5=x4;
    b5=S1(i,2)*x4+S1(i,3);
end
if (y3-S1(i,3))/S1(i,2)>=x3 && (y3-S1(i,3))/S1(i,2)<x4
    a6=(y3-S1(i,3))/S1(i,2);
    b6=y3;
elseif (S1(i,2)*x3+S1(i,3))>y3 && (S1(i,2)*x3+S1(i,3))<y4
    a6=x3;
    b6=S1(i,2)*x3+S1(i,3);
end
if (y4-S1(j,3))/S1(j,2)>x3 && (y4-S1(j,3))/S1(j,2)<=x4
    a7=(y4-S1(j,3))/S1(j,2);
    b7=y4;
elseif (S1(j,2)*x4+S1(j,3))>y3 && (S1(j,2)*x4+S1(j,3))<y4
    a7=x4;
    b7=S1(j,2)*x4+S1(j,3);
end
if (y3-S1(j,3))/S1(j,2)>=x3 && (y3-S1(j,3))/S1(j,2)<x4
    a8=(y3-S1(j,3))/S1(j,2);
    b8=y3;
elseif (S1(j,2)*x3+S1(j,3))>y3 && (S1(j,2)*x3+S1(j,3))<y4
    a8=x3;
    b8=S1(j,2)*x3+S1(j,3);
end
if a5>a7 || b5<b7
    t=a7;a7=a5;a5=t;
    t=b7;b7=b5;b5=t;
end
if a6>a8 || b6<b8

```

```

        t=a8;a8=a6;a6=t;
        t=b8;b8=b6;b6=t;

    end
    k1=(b5-b6)/(a5-a6);k2=(b7-b6)/(a7-a6);k3=(b7-b8)/(a7-a8);
    s2=(k1-k2)*((a5-a6)*(a7-a6)+(b5-b6)*(b7-b6))/(2*(1+k1*k2))+
    (k3-k2)*((a6-a7)*(a8-a7)+(b6-b7)*(b8-b7))/(2*(1+k2*k3));
    s0=s2;
    if s0<s
        s=s0;
        m1=i;
        m2=j;
    end
    if s0<0
        1000000000000000000
    end
end
end
end
%a=S1(m1,1)
%b=S1(m2,1)
m1
m2
if S1(m1,2)==S1(m2,2)
    S1(m1,3)=(S1(m1,3)+S1(m2,3))/2;
    S1(m2,2)=0;
else
    x=(S1(m2,3)-S1(m1,3))/(S1(m1,2)-S1(m2,2));
    y=S1(m1,2)*x+S1(m1,3);
    S1(m1,2)=(S1(m1,2)*(1+S1(m2,2)^2)^0.5+S1(m2,2)*
    (1+S1(m1,2)^2)^0.5)/((1+S1(m2,2)^2)^0.5+(1+S1(m1,2)^2)^0.5);
    S1(m1,3)=y-S1(m1,2)*x;
    S1(m2,2)=0;
end
end
end

```