

中国
青年
科学技术论文精选

中国科学技术出版社

中国青年科学技术论文精选

主 编 周衡孟 严余松 曾光明

副主编 金 磊 王维平 吴唤群

主 办

上 海 交 通 大 学

西 安 交 通 大 学

西 南 交 通 大 学

国 防 科 学 技 术 大 学

湖 南 大 学

湖 南 省 科 学 技 术 协 会

中国科学技术出版社

· 北京 ·

(京)新登字 175 号

图书在版编目(CIP)数据

中国青年科学技术论文精选/周衡孟等编. —北京:中

国科学技术出版社, 1994 • 12

ISBN 7-5046-1877-2

I. 中… II. 周… III. 科学技术—文集 IV. N53

中国版本图书馆 CIP 数据核字(94)第 10945 号

中国科学技术出版社出版发行
北京海淀区白石桥路 32 号 邮政编码:100081
湖南省湘中地质印刷厂印刷

开本:787×1092 毫米 1/16 印张:86 插页:2 字数:214 万字
1994 年 11 月第 1 版 1994 年 11 月第 1 次印刷
印数:1—2000 册 定价:80.00 元

顾 问

(以姓氏笔画为序)

刘 豹	许国志	张钟俊	汪应洛	汪 浩
陈国达	陈启智	陈火旺	李海珍	姚开泰
俞汝勤	袁隆平	顾培亮		

编 辑 委 员 会

(以姓氏笔画为序)

主任委员	俞汝勤			
副主任委员	胡景惠	朱孟明	赵立华	沙基昌
马国馨	王学仁	王浣尘	王柯敏	邓聚龙
宁津生	阳含熙	向元望	孙一坚	刘裔宏
杜 文	李德仁	李怀祖	李新海	朱孟明
吴占松	吴百诗	肖 兵	陈君楷	沙基昌
余铁桥	张新时	张晨曦	张钧华	周宏灏
金宝玺	胡毓达	胡景惠	钟 宏	赵立华
康思奇	曾国雄	谭受清	徐仲林	徐成斌
徐兆月	席裕庚	寇纪淞	顾廷富	

执 行 委 员 会

(以姓氏笔画为序)

王维平	朱克俭	吴少平	吴 林	吴冲锋
吴 敏	吴唤群	张天晓	张 维	张国强
张序君	张鸿雁	陈烈臣	陈宏民	严余松
李敏强	金 磊	周衡孟	姚力建	夏洪胜
徐玖平	席酉民	黄一九	曾光明	曾孔云
黄烟波	潘小松			

目 录

第一部分 理 学

向量约束的半马氏决策规划	刘 克 刘建庸 (1)
两指标随机过程的最优停止的构造	罗建书 (8)
多元位置的 PP 型度量与估计	崔恒建 (15)
一类特殊的非线性最小二乘问题的 ABS 求解方法	王宇平等 3 人 (21)
约束最小二乘的局部影响分析	任仕泉 石 磊 (25)
提高回归曲线模型的分析方法	柳 进等 3 人 (31)
数值输入模糊输出的二次曲线回归模型	许若宁 (35)
广义压缩估计的优良性讨论	张文文 曹树华 (42)
关于参数估计的 Pitman 准则研究	杨 虎 (46)
一类具有良好随机性能的二元周期序列	徐 敏 (53)
GPC 算法的统一格式	刘晓华 (59)
3—状态无向设备网络可靠度计算的分解定理	李东魁 迟汉忠 (62)
零曲率表示与特征算子方程	马文秀 (68)
不含求逆运算的换元法	李志斌 邵 燕 (76)
一类分段常数混合自变量方程的振动性	刘树堂 (81)
带权 P—Laplace 算子第一特征值的性质	张子方 (89)
求解非线性变分不等式的线性化方法的收敛性研究	李董辉 曾金平 (94)
多重迭代波算子的初值基本解	同登科 (99)
三阶非线性微分方程的周期边值问题	王国灿 (107)
Volterra 型积分方程配置解的加速收敛分析	胡齐芽 (111)
高阶非线性中立型方程正解的存在性	周 勇 (117)
一类一阶泛函微分方程解的存在性与分支	时 宝 (125)
奇异离散系统的延展稳定性	温香彩 刘永清 (130)
V 函数法在测度微分大系统稳定性中的应用	关治洪 刘永清 (135)
线性系统稳定度的实用准则	黄廷祝 (142)
关于离散时变线性系统的标准分解	吕 靖 曹艳萍 (150)
多变量有限元 P—收敛的事后误差估计	段 梅 周本宽 (156)
一类数值求解 $y''=f(x,y)$ 具有极小相位延迟误差的 4 阶 Noumerov 型显式和隐式法	向开理 (159)
广义系统的极点配置问题	葛照强 (165)

一个求解二次规划总体最优解的新方法	曾安军等 3 人 (172)
求解线性约束凸规划的一种分解方法	谭中富 张作泉 (177)
线性分式规划的多项式时间算法	简金宝 蓝 敏 (181)
非线性规划共轭投影梯度法及其收敛性	时贞军 (186)
求解凸规划问题的递缩约束集算法	秦志林 吴冲锋 (193)
解线性约束最优化问题的一个新算法	韦建芳 (199)
一类模糊终端约束的资源分配问题	任 军 (207)
共形可压缩区域上的一个逼近定理	马继钢 何育赞 (211)
S—连续映射注记	陈仪香 (218)
具有紧 k —网的空间	林 寿 李进金 (222)
有限单群的一种数量刻划	陈贵云 (227)
被可解子群特征化的非可解群	黎先华 (236)
TL 子群和 Ω —TL 子群	王住登 (244)
Fuzzy 子环的商环与直和	张诚一 (251)
指数函数级数表示的分形的盒子维数与分数阶导数	袁苗龙 邓冠铁 (256)
一种检验 Z 域系统稳定的新方法 Z 域正实函数的应用	毕卫红等 3 人 (260)
Chrestenson 函数与广义桥函数	竺南直 (263)
函数拟凸的充要条件	杜 江 (269)
关于集值函数的微分	元昌安 (272)
模糊值函数的微积分中值定理	党发宁 (278)
微积分与近代数学	黄书亭 (284)
开曲线的凸性研究	童照春 王兴波 (292)
平面 B ——样条曲线的拐点定理	王兴波 童照春 (296)
与给定多边形相切的样条逼近曲线	方 遼 王文仲 (301)
一类四角仙人掌图的优美性	李秀芬 路 线 (304)
图 K —优美性的必要条件	王金超 (308)
交通规划中的一个最小费用流模型的改进算法	吴唤群 唐 莉 (311)
有限输入输出加权模式的标量型谱实现	唐万生等 3 人 (316)
最优火力分配与相对作战指数的确定	张序君 沙基昌 (320)
功能型修复不如新系统广义可靠性分析	马云东 孙宝铮 (325)
具有投资滞后动态投入产出模型及其最优控制	李少文 刘晓华 (331)
灰色投入产出模型的位移计算	韩旭里 李松仁 (334)
多准则群决策的一致性测度	方 然 (338)
具有小干扰的控制系统问题及其微分对策模型	邵 卫 (342)
群决策问题的一种求解方法	武小悦 (348)
一个单周期多品种存货模型与基于边际分析理论的算法	易 发 (351)
S—ALOHA 系统中平均周期计算方法研究	赵东风等 4 人 (354)
关于情景分析中消除非一致性模型的研究	宗蓓华 施 欣 (359)
异质结构界面应力的实验研究	亢一澜 张继平 (364)

承压地层中弹性流体径向流——固耦合渗流问题的摄动解	徐曾和 刘立忱 (369)
无坐力炮不平衡力的识别法	李毓华 戈新民 (375)
拱坝塑性极限分析	范金星 (379)
水平排渗体参数对渗流自由面型态的影响	沈振中 速宝玉 (384)
飞行器增升效应的若干新方法研究	李 峰 (389)
托卡马克边缘湍流中的欧姆耗散	胡双辉等 3 人 (398)
一类移位系统的复杂性态	傅新楚 (403)
玻色子有效相互作用与四玻色子核的核谱	冯 芒 (411)
量子力学的信息论表述方法	师志恒 (416)
随机信号的非线性变换及频谱响应分析	沈民奋 孙丽莎 (420)
水下声强的测量与应用研究	杨德森 (425)
N^2 递推 Y 矩阵算法求 N 体问题的多散射解	任 伟等 4 人 (433)
直流电路中自由电荷与稳恒电流的关系	周一平 (439)
$YBa_2Cu_3O_{7-x}$ 外延膜平面电阻率的各向异性	程香爱等 4 人 (444)
热处理对非晶态 Fe—Si—B 合金的磁性影响研究	王玲玲等 4 人 (448)
粉尘扩散系数的实用预报方法	章社生 (452)
高模量 PAN 碳纤维 X—射线结构分析	苏玉长等 5 人 (457)
离子选择电极阵列广义标准加入法同时测定伯、仲、叔胺	王柯敏等 3 人 (461)
丁苯橡胶对水泥石物理力学性能的改善及对水泥水化过程的影响	梁乃兴 (466)
基于强疏水性荧光素衍生物的荧光增强的羧酸光纤化学传感器的研究	曾惠惠等 4 人 (471)
液体的蒸发热与 Anotine 常数的换算	文立中 田德余 (476)
苯乙烯—4—(甲基丙烯酸)2,2,6,6—四甲基哌啶醇酯共聚物 (P. D. S)	
对聚苯乙烯光稳定化研究	刘晓暄等 3 人 (482)
聚苯乙烯光氧化行为的早期表征及其光谱特征研究	刘晓暄等 5 人 (487)
温度对萃取回收水中醋酸的影响研究	杨春平等 5 人 (492)
酸性废水的微生物处理实验研究	王恩德 于志凯 (495)
真菌与龙血树血竭形成关系的初步研究	王兴红等 4 人 (500)
米曲霉 (<i>Aspergillus Oryzae</i>) 菌种的研究	向 洋等 3 人 (504)
火腿生香酵母 YN8662 菌株的鉴定及其风味化合物的分析	王登全等 4 人 (509)

第二部分 工 学

安徽铜陵地区基底深断裂及其控矿机理	毕 华 (513)
陕西紫阳富硒茶的富硒规律及其地质特征	雒昆利 (519)
东濮凹陷东南部下第三系盆地分析及油气评价	朱筱敏等 4 人 (525)
阳新统油气储集和渗流结构的分维研究	汪富泉 (535)
放顶煤综采覆岩“两带”高度分布规律研究	麻凤海等 4 人 (539)
地表下介质横向非均匀对埋置管道地震反应的影响	袁晓铭等 3 人 (544)

黄土地区的地震灾害预测	王兰民 (553)
地面沉降预测计算的 SA 模型	桂和荣等 4 人 (560)
泥石流——防治系统及其模式的初步研究	唐晓春 (566)
泥石流防治模式初探	唐晓春 (570)
航空遥感技术在立地质量评价中的应用研究	谭俊 (579)
矿压实测数据缺失的最佳拟补方法	谭允亮等 3 人 (583)
生物材料用钛及其合金表面氧化物覆膜改性研究	黄楠等 3 人 (587)
铅、锌(银)、多金属矿床矿石类型定量划分的方法	周殿铭等 4 人 (592)
铜合金中微量锡的测定	王萍 (598)
低合金抗磨铸铁中碳化物分析与基体组织选择	陈有志 (601)
熔融盐结构性质的计算机模拟研究	谢刚 宋宁 (606)
矿仓组合放矿控制专家决策支持系统	舒航 (611)
矿山井巷工程神经网络专家系统的理论与实践	魏一鸣 (616)
两足步行机器人前向动态步行综合	马宏绪等 3 人 (621)
五节臂伸缩机构的创新设计	金霞等 3 人 (628)
磁性传感汽车转向自动指示器的设计	李保谦等 3 人 (632)
静电弥雾机的设计研究	赵朝会等 4 人 (635)
电力机车微机测控系统的设计研究	郭其一等 3 人 (639)
农机化管理信息系统的研制与实现	何勇等 3 人 (645)
采用新型控制方案的 CSIM 微机控制系统	王静 (650)
抽油机井系统效率预测与参数优选的新方法	董世民 姚春冬 (655)
OTA—C HIGHPASS LADDER FILTER	Wu Jie (660)
计算有功静态安全域的新扩展算法	朱继忠 (664)
求多机电力系统特征值的预估—校正算法	谷根代等 3 人 (669)
多分区混合电力系统规划数学模型研究	赵敏 (673)
长湖水电站尾水河水力特性的研究	赖翼峰 (680)
机电一体化技术学科理论与体系结构	潘仲明 (685)
面向对象的多媒体仿真技术:一个研究原型	王维平 朱一凡 (690)
基于变换的软件开发方法及其支持环境	殷建平 (695)
解释语言调试工具的功能、设计与实现	马小宏 (702)
建造避碰专家系统方法初探	翁跃宗 (706)
决策支持系统中知识库系统的一种设计方法	吴启明 (710)
一种新型的专家系统——可拓专家系统	李健 王行愚 (715)
公司决策支持系统 (CDSS) 设计	潘胜强 (721)
在用工业建筑物动态管理与决策专家系统	杨新华 (727)
中高层砌体结构房屋设计——介绍一种新的结构体系	刘迎五 梁建国 (731)
灰色系统关联分析在塔板评价与选型中的应用	张长银 朱家才 (735)
大型人字门设计计算中的若干问题	张燎军 傅作新 (740)
房间空调器的模糊控制方法	张国强等 4 人 (746)

聚氯乙烯给水管的开发与应用技术研究	韩洪军等 3 人 (750)
建筑工程 AHP 评标法软件系统的研究	邓铁军等 3 人 (755)
惯性力法随机有限元重力坝动力可靠度全概率分析	孙长龙 吕泰仁 (758)
货物运输结构演变自组织的机理分析	赵一平 (764)
城市公交线网优化模型的改进及程序设计	晏启鹏 (769)
双线自动闭塞区段组合列车有关问题的探讨	严余松 (773)
矿石进长江运输系统研究——模型及其应用	赵 刚 (777)
多因素影响程度时变的港口吞吐量预测方法的研究	兰培真 (782)
城市污水处理厂系统最优规划研究	梅亚东 (786)
城市多种污染源下浅层地下水水质的定量预测方法及其应用	桂和荣等 4 人 (792)
中国渤海莱州湾滨海地区海水入侵灾害形成条件的初步分析	季明川 (798)
黄河口北部海域沉积的氧化还原环境研究	宋金明等 7 人 (803)
环渤海地区主要自然灾害与防御对策	高 红 (811)
河流水质灰色模型中灰参数的最优估计	曾光明等 5 人 (816)
环境系统规划的灰色方法	曾光明等 5 人 (823)
河流水质系统的灰色模拟计算方法	曾光明等 5 人 (831)
湖南凌津滩库区造纸废水对水质的影响及治理措施研究(上)	曾光明等 6 人 (839)
湖南凌津滩库区造纸废水对水质的影响及治理措施研究(下)	曾光明等 6 人 (844)
灰色建模的置零法则及其在噪声影响研究中的应用	郭 原 郭文扬 (850)

第三部分 农学与林学

节水型农业水管理中几个基础理论问题的探索	康绍忠 (855)
黄土高原水土资源组合特征与调控研究	赵艺学 (860)
湖泊水库富营养化程度模糊评价模式	熊德琪 (866)
以离异差平方和为判据进行群落排表分类	张金屯 (873)
改土培肥是农业持续稳定发展的必由之路	邓福军 (877)
海南岛农业生态环境质量综合评判	周兆德 (881)
商丘地区主要气候条件对小麦产量的影响及对策	王建设等 8 人 (890)
豫东黄潮土区小麦施钾技术效应研究及推广应用进展	张慎举等 14 人 (894)
湘西晒红烟品种产量潜力的遗传分析和新品种选育	李国民 王复文 (901)
环境温度对水稻光敏雄性不育性遗传分离影响的初步研究	程式华等 3 人 (904)
温度对光照诱导光敏感核不育水稻的发育与育性转变的影响初步研究	贺浩华等 3 人 (908)
稻米品质多性状同步间接选择的数学模型探讨	唐启源等 3 人 (913)
张掖市新开荒地水稻产量模型的研究	徐存才 (918)
水稻土上小麦锰素营养的研究	吕世华 (921)
黄红紫泥小麦生产系统肥料效应分析与调控	魏世强等 3 人 (927)
核质杂种小麦的育性分析及新型雄性不育系的研究	易自力等 3 人 (932)

玉米细胞质雄性不育性的激素调控机制	夏 涛 刘纪麟 (937)
中国野生大豆 (G. soja) 种子贮藏蛋白的电泳分析	庄炳昌等 5 人 (942)
黄淮平原大豆品种的主要数量性状对产量稳定性影响的分析	刘 辉等 4 人 (946)
夏大豆发生“荚而不实”的生态技术条件及预防措施	侯乐新等 10 人 (949)
论葡萄酒生产系统——兼论水平分科的必要性	李 华 王 华 (953)
园艺作物化肥应用的现状及展望	王 倩等 3 人 (958)
果树营养及其诊断	韩振海等 3 人 (963)
桃李快速育苗技术研究的初步探讨	王仁才 (967)
国营农场发展“二高一优”农业的思考	武 志 柳永芬 (971)
害虫对菊酯类农药的抗性机制及应用	戴小枫 韩丽娟 (974)
虹鳟甲状腺对大马哈鱼垂体激素反应特异性的研究	陈松林 B. Breton (980)
鸡粪养猪配套技术的研究	曾勇庆等 12 人 (986)
桑蚕微粒子病集团母蛾检验方法研究	徐世清 王林芳 (993)
张家界国家森林公园风景质量评价及建设对策	李世东 邓金阳 (1003)
贡嘎山东坡森林土壤有机质的垂直分布规律	王良健 (1009)
陇南木本植物资源调查研究	柴发熹 (1014)
松嫩平原中部植物群落的间接梯度分析与数量分类	孙国荣等 4 人 (1021)
卷柏配子体发育及其制约因素的研究	刘保东 孙国荣 (1028)
卷柏科孢子囊排布规律的系统学意义初步研究	刘保东等 4 人 (1033)
两种卷柏幼孢子体形态建成的系统学研究	刘保东等 3 人 (1038)

第四部分 医 学

新型抗肿瘤生物“导弹”——胃癌特异性免疫脂质体的研究	徐 梁等 5 人 (1045)
彩色多普勒联合脉冲多普勒诊断小肝癌的研究	王文平等 4 人 (1050)
苯乙酮及其衍生物缩氨基硫脲与铜配合物的抗癌活性	赵喜荣等 3 人 (1054)
重组人肿瘤坏死因子抗卵巢癌的特性	赵恩峰等 8 人 (1058)
无症状心肌缺血发病机制及临床的研究	吴 林 (1063)
特异性 i——RNA 治疗前后婴儿 NK IFN 和 IL——2 活性变化的研究	卓安山等 5 人 (1067)
临床用药性别化的药理学基础	谢红光 (1072)
肺微循环内血液流动的模式	李 丽 葛照强 (1074)
四环素肾盂灌注治疗乳糜尿	杨金瑞等 5 人 (1082)
三叉宁治疗原发性三叉神经痛的实验研究	曲晓璐等 3 人 (1085)
下肢全息穴区穴位分布规律新探	戴 会 白 纯 (1088)
肝豆状核变性培养细胞内铜、金属硫蛋白代谢的综合研究	陈 嵘等 4 人 (1092)
旋毛虫新生幼虫对小鼠的免疫保护作用研究	朱兴全等 5 人 (1098)
绞股蓝总皂甙对体外培养小鼠脾淋巴细胞增殖的影响与 DNA 合成的关系	廖端芳等 5 人 (1101)

超声成像系统声像图几何失真分析与校正方法的研究	罗建	卞正中 (1104)
农村健康保险的精算体系	李良军	杨树勤 (1109)

第五部分 交叉科学与其他

FDYPECM 研究——动态模式经济控制论模型的模糊形式研究		
.....	王意冈	王浣尘 (1115)
简略寿命表编制的数理技术与实例		张鸿雁 (1121)
人力系统与子系统	王国秋	沙基昌 (1126)
试论灾害系统观及其灾害系统学的研究方法	李本海	赵纯均 (1138)
安全减灾系统工程原理及评价模型论		金磊 (1143)
信息资源管理与数据库产业		路节 (1151)
信息生产率测度的数学模型		刘廷元 (1156)
战术导弹寿命周期费用的不确定性灰模处理		章敬东 (1161)
GAM 模型在导弹投资预测中的应用		苏洪潮 (1169)
科技书刊出版与科技进步的系统分析	周衡孟	潘胜强 (1172)
企业经营管理的系统分析与模拟研究		张晓 (1177)
企业逆境管理理论的原理与应用		余廉 (1183)
国际性企业现金管理模型		程仕军 (1189)
大型工程项目建设的风险及风险管理	周直	何秀杰 (1194)
多品种、小批量生产作业计划管理系统的开发和实施		潘郁 (1198)
一种企业生产决策的新方法		刁兆峰 (1203)
投资系统多个互斥方案评价准则的探讨	刘平	郭宏 (1208)
房地产价格收益还原评估法中还原利率的确定		施建刚 (1211)
产业规划决策支持系统研究	高麟等 3 人	(1214)
结构变化与经济增长的定量分析法		张朝敏 (1218)
产量—质量—收益综合计酬法探讨		邱国栋 (1224)
产品设计的参数优化方法研究	单汨源	王艾伦 (1229)
大规模地面——地下水资源联合运用系统管理模式研究	朱文彬	闫瑞玲 (1235)
菜蓝子工程灰色系统仿真模型		尹琦 (1242)
市场场系统论		马培荣 (1246)
市场经济条件下科技人员政策建议	孙萍	张立法 (1251)
信息产业人才的国际流动及我国的挑战与对策	刘廷元等 4 人	(1258)
我国金融信托业发展的思考		王江平 (1263)
市场经济条件下煤炭企业级差收支调节政策研究	周德群	柏长志 (1267)
关于科学技术是第一生产力的再认识		郭庆存 (1270)
中国科学技术进步与社会控制		陈光 (1276)
试论科技、经济、社会发展中的“循环加速器”	刘国涛	帅相志 (1281)
评价地域科技能力的一种新思路新方法		周根贵 (1284)
区域经济自组织机制		刘思峰 (1289)

R&D项目评估活动的模式	童 健 连燕华 (1293)
省属科研机构综合评价方法	田乐远等 3 人 (1299)
中试基地面临的困境及对策研究	姚先桥 (1303)
组配促需:资源效用型林业发展的机制与政策	李智勇 (1307)
平湖市工业发展优化模拟	施 放等 4 人 (1313)
试论英语修辞研究的三位一体取向	申屠菁 赵国浩 (1320)
灰色关联分析用于课程设置体系的整体优化	吴 强等 4 人 (1325)
体育专业田径普修课考核内容与办法初探	马尉平 严 秋 (1328)
我国大学生体育生活方式的社会学探析	荆光辉等 3 人 (1332)
青少年儿童体质评价方法新探讨	江 兵 洪 雷 (1338)
向后类高难度空翻起跳技术的研究	李艳翎 戈炳珠 (1342)
服装设计师自身素质与设计定位	辛艺华 (1347)

向量约束的半马氏决策规划*

刘克 刘建庸

(中科院应用数学研究所, 北京 100080)

【提要】 本文讨论了具有向量约束的折扣半马氏决策规划(SMDP), 同时给出了这一模型的算法。

【关键词】 半马氏决策规划(SMDP) 向量约束 最优策略

1 引言

到目前为止, 带有向量约束的 SMDP 尚未见到有人讨论。1985 年, 在参考文献[1]中讨论了具有约束的平均马氏决策(MDP), 但是参考文献[1]并未给出有效的算法。本文首次讨论具有向量约束的 SMDP。第二节中我们先给出模型; 第三节讨论了有限跳跃模型并给出了该模型的算法和删除规则; 第四节给出了对具有向量值约束的 SMDP 模型的带有删除规则的算法。

2 模型

一个有约束的 SMDP 为一八重组 $\{S, A, q, t, r, c, \alpha, v\}$, 其中, S 为状态空间 A 为行动集两者皆为有限集合; q 为时齐的状态跳跃簇; t 为状态跳跃的时间分布; r 为报酬函数; c 为费用函数; $\alpha > 0$ 为折扣因子; $v \in R^{|S|}$ 为一向量约束, 其中 $|S|$ 表示 S 中元素的个数。

映射 $f: S \rightarrow A$ 为决策函数。决策函数全体记为 F 。令 $f_i \in F, i=0, 1, 2, \dots (f_0, f_1, \dots) = \pi$ 称为马氏策略, 其全体记为 Π 。令 $f \in F, (f, f, \dots) = f^\infty$ 称为平稳策略, 其全体记为 Π_∞ 。

当系统从状态 $i \in S$ 出发, 使用策略 $\pi \in \Pi$, 其期望折扣总报酬和总费用为

$$R(i, \pi) = \int_0^\infty e^{-\alpha T} R(dT | \pi, i) \quad (1)$$

$$C(i, \pi) = \int_0^\infty e^{-\alpha T} C(dT | \pi, i) \quad (2)$$

其中, $R(T | \pi, i)$ 为从 $t=0$ 到 T 的期望总报酬, $C(T | \pi, i)$ 为期望总费用, 均是在使用策略 π 并从状态 i 出发的条件下。并用 $R(T | \pi)$ 和 $C(T | \pi)$ 记为它们的列向量。

$$\forall a \in A, i \in S, \text{ 令 } \beta_a(i, a) = \int_0^\infty e^{-\alpha u} t(du | i, a) \quad (3)$$

假设 1

$$\beta_i = \max_{\substack{a \in A \\ i \in S}} \beta_a(i, a) < 1 \quad (4)$$

* 云南省青年基金资助项目

作者简介 刘克, 男, 1956 年 7 月出生。1982 年夏毕业于中国科技大学数学系运筹学专业, 获理学学士学位。现任副研究员。发表论文 20 余篇。



令向量 $u_1, u_2 \in R^p$, $u_1(i)$ 和 $u_2(i)$ 分别记其第 i 个分量, 若 $u_1(i) \leq u_2(i)$ (或 $u_1(i) < u_2(i)$), $i = \overline{1, p}$, 则记为 $u_1 \leq u_2$ ($u_1 < u_2$)。若 $u_1 \leq u_2$ 不成立, 记为 $u_1 \not\leq u_2$, 同样定义 $u_1 \prec u_2$ 。

令 $D = \{\pi \in \Pi | C(\pi) \leq v\}$, 其中 $C(\pi)$ 的第 i 个分量是 $C(i, \pi)$ 。

定义 2.1 令 $(\epsilon, \epsilon') \geq 0$, $\pi^* \in \Pi$ 。若 $C(\pi^*) \leq v + \epsilon' \mathbf{1}$ 且 $\forall \pi \in \Pi$ 有 $R(\pi^*) \geq R(\pi) - \epsilon \mathbf{1}$, 则 π^* 为 (ϵ, ϵ') -最优策略 (对于带约束的 SMDP), 其中 $R(\pi)$ 的第 i 个分量为 $R(i, \pi)$, $\mathbf{1}$ 为分量均为 1 的列向量。一个 $(0, 0)$ -最优策略简称为最优策略。

注: 约束 SMDP 可能不存在 Π 中的最优策略。

由于 SA 有限, 不失一般性, 我们可设 $0 \leq r(i, a) \leq M_1, 0 \leq c(i, a) \leq M_2, i \in S$ 。

引理 2.1 $\forall \pi = (f_0, f_1, \dots) \in \Pi, i \in S$, 有

$$\begin{aligned} R(i, \pi) &= \int_0^\infty e^{-\alpha T} R(dT | \pi, i) \\ &= r(i, f_0) + \beta_a(i, f_0) \sum_{i_1} q(i_1 | i, f_0) r(i_1, f_1) \\ &\quad + \beta_a(i, f_0) \sum_{i_1} q(i_1 | i, f_0) \beta_a(i_1, f_1) \sum_{i_2} q(i_2 | i_1, f_1) r(i_2, f_2) + \dots \\ &\triangleq [f_0]r(i) + [f_0\beta_a q r](i) + [f_0\beta_a q f_1\beta_a q r](i) + \dots \triangleq \sum_{k=0}^\infty [\pi\beta_a q r]_k(i) \end{aligned} \quad (5)$$

$$\text{类似地有 } c(i, \pi) = \sum_{k=0}^\infty [\pi\beta_a q c]_k(i) \quad (6)$$

引理 2.1 的证明可见参考文献[2]的引理 2。有时(5, 6)也记为向量形式:

$$R(\pi) = \sum_{k=0}^\infty [\pi\beta_a q c]_k = [f_0 r] + [f_0\beta_a q R(\pi)] \quad C(\pi) = \sum_{k=0}^\infty [\pi\beta_a q c]_k = [f_0 c] + [f_0\beta_a q C(\pi)]$$

其中, $\pi = (f_0, f_1, \dots)$, ${}^1\pi = (f_1, f_2, \dots)$, 均属于 Π 。

3 有限跳跃模型

有限跳跃模型为 $\{S, A, q, t, r, c, \alpha, v, N\}$, 其中, $S, A, q, t, r, c, \alpha, v$ 均同前; N 均为正整数表示跳跃次数。简记为 FCSMDP。

定义 3.1 $\pi^{(n)} = (f_n, f_{n+1}, \dots, f_N)$ 为一从 n 次跳跃到 N 次跳跃的马氏策略, 简称为 n -策略, 其中 $f_i \in F, i = \overline{n, N}, 0 \leq n \leq N$ 。 $\Pi(n)$ 为全体 n -策略。 只在本节中, 0 -策略 $\pi^{(0)}$ 记为 π 。

对 $\pi \in \Pi(0), \forall i \in S, N$ 跳跃期望报酬和费用定义为

$$R_N(i, \pi) = \sum_{k=0}^N [\pi\beta_a q r]_k(i) \quad (7)$$

$$C_N(i, \pi) = \sum_{k=0}^N [\pi\beta_a q c]_k(i) \quad (8)$$

在第 n 次跳跃系统处于状态 i , 用 $\pi^{(n)}$, 到 N 的期望报酬和费用为

$$\bar{R}_n(i, \pi^{(n)}) = \sum_{k=0}^{N-n} [\pi^{(n)}\beta_a q r]_k(i) \quad (9)$$

$$\bar{C}_n(i, \pi^{(n)}) = \sum_{k=0}^{N-n} [\pi^{(n)}\beta_a q c]_k(i) \quad (10)$$

其中, $i \in S, 0 \leq n \leq N$ 。

令 $U = \{\pi \in \Pi(0) | \bar{C}_0(\pi) = C_N(\pi) \leq v\}$, 其中 $\bar{C}_0(\pi)$ 和 $C_N(\pi)$ 均为向量, 其第 i 个分量分别

为 $\bar{C}_0(i, \pi)$ 和 $C_N(i, \pi)$ 。

定义 3.2 令 $\Gamma \subset \Pi(0)$, $\epsilon \geq 0$, $\pi^* \in U \cap \Gamma$ 。若 $\bar{R}_0(\pi^*) = R_N(\pi^*) \geq \bar{R}_0(\pi) - \epsilon 1 = \bar{R}_N(\pi) - \epsilon 1$, 对于 $\forall \pi \in U \cap \Gamma$, 则 π^* 称为 *FCSMDP* 的 ϵ -最优策略(对 Γ), 其中 $\bar{R}_0(\pi)$ 和 $R_N(\pi)$ 为分量 $\bar{R}_0(i, \pi)$, $R_N(i, \pi)$ 的列向量。

注: 存在例子表明不存在 *FCSMDP* 的 ϵ -最优策略(对 $\Pi(0)$)。

令 $\bar{r}(i, a) = (r(i, a), -C(i, a))^T$, $a \in A, i \in S$ 。则我们可得到向量值 N 跳跃 *SMDP* (记为 *VFSMDP*), 即 $\{S, A, q, t, \bar{r}, \alpha\}$ 。当第 n 次跳跃系统处于状态 i 时, 用 $\pi^{(n)}$ 则期望总报酬为

$$V_n(i, \pi^{(n)}) = \sum_{k=0}^{N-n} [\pi^{(n)} \beta_0 \bar{q} r]_k(i) = (\bar{R}_n(i, \pi^{(n)}) - \bar{C}_n(i, \pi^{(n)}))^T \quad (11)$$

其中, $i \in S, \pi^{(n)} \in \Pi(n), 0 \leq n \leq N$ 。

令 $\pi_1^{(n)}, \pi_2^{(n)} \in \Pi(n)$, 若 $V_n(i, \pi_1^{(n)}) \leq V_n(i, \pi_2^{(n)})$ 对一切 $i \in S$, 则记为 $V_n(\pi_1^{(n)}) \leq V_n(\pi_2^{(n)})$, 若 $V_n(i, \pi_1^{(n)}) = V_n(i, \pi_2^{(n)}), i \in S$, 则记为 $V_n(\pi_1^{(n)}) = V_n(\pi_2^{(n)})$ 。

定义 3.3 令 $0 \leq n \leq N, \Lambda \subset \Pi(n)$ 及 $\pi_0^{(n)} \in \Lambda$, 称 $\pi_0^{(n)}$ 在 Λ 中有效是指: $\forall \pi^{(n)} \in \Lambda$, 若 $V_n(\pi_0^{(n)}) \leq V_n(\pi^{(n)})$, 则有 $V_n(\pi_0^{(n)}) = V_n(\pi^{(n)})$ 。

令 $e\Lambda = \{\pi^{(n)} \in \Lambda | \pi^{(n)} \text{ 在 } \Lambda \text{ 中有效}\}$ 。并令

$$F_n = e(\Pi(N)) \quad F_{K-1} = e(F \times F_K), 1 \leq K \leq N$$

其中, $F \times F_{K-1} = \{(x, y) | x \in F, y \in F_K\}$ 。易见 $F \times F_K \subset \Pi(K-1), 1 \leq K \leq N$ 。由于 F 有限, 故 $F \times F_K (1 \leq K \leq N)$ 有限。于是可找到每个有效集 $e(F \times F_K)$ 。

引理 3.1 令 $0 \leq n \leq N, \Lambda \subset \Pi(n), \forall \pi^{(0)} \in \Lambda$, 存在 $\tilde{\pi}^{(n)} \in \Lambda$ 满足 $V_n(\tilde{\pi}^{(n)}) \leq V_n(\pi^{(n)})$ 。

证明 由 Λ 的有限性, 证明显然。

定理 3.1 (i) $F_n \subset e(\Pi(n)), 0 \leq n \leq N$ 。

(ii) 令 $0 \leq n \leq N, \forall \pi^{(n)} \in \Pi(n)$, 存在 $\tilde{\pi}^{(n)} \in F_n$, 使得 $V_n(\pi^{(n)}) \leq V_n(\tilde{\pi}^{(n)})$ 。

证明 (用归纳法) 由定义 $F_n = e(\Pi_N) \subset e(\Pi_N)$ 。 $\forall \pi^{(N)} \in \Pi(N)$, 由引理 3.1, $\exists \tilde{\pi}^{(N)} \in e(\Pi(N)) = F_N$, 满足 $V_N(\pi^{(N)}) \leq V_N(\tilde{\pi}^{(N)})$, 即时 $n=N$ 成立。

设当 n 时成立 ($0 \leq n \leq N$)。

$\forall \pi^{(n-1)} \in F_{n-1}$, 令 $\pi_1^{(n-1)} \in \Pi(n-1)$ 及 $V_{n-1}(\pi^{(n-1)}) \leq V_{n-1}(\pi_1^{(n-1)})$ 。令 $\pi_1^{(n-1)} = (f, \pi_1^{(n)})$ 。由归纳假设, $\exists \tilde{\pi}_1^{(N)} \in F_n$ 满足 $V_n(\pi_1^{(n)}) \leq V_n(\tilde{\pi}_1^{(n)})$ 。

记 $\tilde{\pi}_1^{(n-1)} = (f, \tilde{\pi}_1^{(n)})$, 易见

$$\begin{aligned} V_{n-1}(i, \pi_1^{(n-1)}) &= \bar{r}(i, f(i)) + \beta_n(i, f(i)) \sum_j q(j|i, f(i)) V_n(j, \pi_1^{(n)}) \\ &\leq \bar{r}(i, f(i)) + \beta_n(i, f(i)) \sum_j q(j|i, f(i)) V_n(j, \tilde{\pi}_1^{(n)}) = V_{n-1}(i, \tilde{\pi}_1^{(n-1)}) \quad i \in S \end{aligned}$$

由于 $\tilde{\pi}_1^{(n-1)} \in F \times F_n$ 且 $\pi^{(n-1)} \in F_{n-1} = e(F \times F_n)$, 于是有 $V_{n-1}(\pi^{(n-1)}) = V_{n-1}(\tilde{\pi}_1^{(n-1)}) = V_{n-1}(\pi_1^{(n-1)})$ 。故 $\pi^{(n-1)} \in e(\Pi(n-1))$, 即 $F \subset e(\Pi)$ 。于是 (i) 对 $n-1$ 成立。

$\forall \tilde{\pi}^{(n-1)} \in \Pi(n-1)$ 由上面的证明, $\exists \tilde{\pi}^{(n-1)} \in F \times F_n$, 满足 $V_{n-1}(\tilde{\pi}^{(n-1)}) \leq V_{n-1}(\tilde{\pi}_1^{(n)})$, 由引理 3.1, $\exists \tilde{\pi}_2^{(n-1)} \in e(F \times F_n)$, 满足 $V_{n-1}(\tilde{\pi}_1^{(n-1)}) \leq V_{n-1}(\tilde{\pi}_2^{(n)})$ 。所以 $V_{n-1}(\pi^{(n-1)}) \leq V_{n-1}(\tilde{\pi}_2^{(n)})$, 即 (ii) 对 $n-1$ 成立。定理证毕。

令 $W_1 = \{\pi \in F_0 | \pi \text{ 为 } FCSMDP \text{ 的 } 0\text{-最优策略(对 } F_0)\}$ 。

$\varepsilon > 0, W_2 = \{\pi \in F_0 \mid \pi \text{ 为 } FCSMDP \text{ 的 } \varepsilon\text{-最优策略 (对 } F_0)\}$ 。

定理 3.2 (i) 若 $\pi \in W_1$, 则 π 为 $FCSMDP$ 的最优策略 (对 $\Pi(0)$); 若 $W_1 = \emptyset$, 则 $FCSMDP$ 没有最优策略 (对 $\Pi(0)$)。

(ii) 若 $\pi \in W_2$, 则 π 为 $FCSMDP$ 的 ε -最优策略 (对 $\Pi(0)$); 若 $W_2 = \emptyset$, 则 $FCSMDP$ 没有 ε -最优策略 (对 $\Pi(0)$)。

(iii) 若 $\{\pi \in F_0 \mid \bar{C}_0(\pi) \leq v\} = \emptyset$, 则 $U = \emptyset$ 。

证明略。

定理 3.2 实际上是第四节算法的一部分, 为使算法更为有效, 我们引入删除 $F_n (0 \leq n \leq N)$ 中一些不必要元素的规则。

令 $v^* = \max_{j \in S} v_j$, 其中 v_j 为向量 V 的第 j 个分量。又令 $\beta_1 = \max_{\substack{a \in A \\ j \in S}} \beta_2(i, a)$ 。

假设 2 $\beta_1 > 0$

在假设 2 下, 删除规则如下:

令 $\bar{F}_N = F_N - \{\pi^{(N)} \in F_N \mid \bar{C}_N(\pi^{(N)}) > \frac{v^*}{\beta_1^N}\}$

$\bar{F}_{k-1} = e(F \times \bar{F}_k) - \{\pi^{(k-1)} \in e(F \times \bar{F}_k) \mid \bar{C}_{k-1}(\pi^{(k-1)}) > \frac{v^*}{\beta_1^{k-1}}\}$ 其中 $0 \leq k \leq N$ 。

定理 3.3

(i) 若 $V \neq \emptyset$, 则 $\bar{F}_n \neq \emptyset, 0 \leq n \leq N$ 。

(ii) $\bar{F}_n \subset F_n, 0 \leq n \leq N$ 。

(iii) 令 $0 \leq n \leq N, \forall \pi^{(n)} \in \Pi(n)$ 。若 $\bar{C}_n(\pi^{(n)}) \not\geq \frac{v^*}{\beta_1^n}$, 则 $\exists \tilde{\pi}_n^{(n)} \in \bar{F}_n$, 满足 $V_n(\pi^{(n)}) \leq V_n(\tilde{\pi}_n^{(n)})$ 。

证明 (用向后归纳法) 由定义 $F_N \subset F_N$ 。 $\forall \pi^{(N)} \in \Pi(N)$ 令 $\bar{C}_N(\pi^{(N)}) \not\geq \frac{v^*}{\beta_1^N}$ 由定理 3.1, $\exists \pi^{(N)} \in F_N$, 满足 $V_N(\pi^{(N)}) \leq V_N(\tilde{\pi}^{(N)})$ 。由 (11) $\bar{C}_N(\pi^{(N)}) \geq \bar{C}_N(\tilde{\pi}^{(N)})$ 故 $\bar{C}_N(\tilde{\pi}^{(N)}) \not\geq \frac{v^*}{\beta_1^N}$, 即 $\tilde{\pi}^{(N)} \in \bar{F}_N$ 。

令 $\pi = (f_0, f_1, \dots, f_{N-1}, \pi^{(N)}) \in \Pi(0)$ 且 $\bar{C}_0(\pi) \leq v$, 因为 $c(i, a) \geq 0, a \in A, i \in S$, 易见

$$\bar{C}_0(\pi^{(N)}) = \frac{N}{k=0} [\pi \beta_a q c]_k \geq [\pi \beta_a q c]_N$$

若 $\bar{C}_N(\pi^{(N)}) \geq \frac{v^*}{\beta_1^N}$, 我们知 $\bar{C}_0(\pi) > v^*$, 矛盾。故 $\bar{C}_N(\pi^{(N)}) \not\geq \frac{v^*}{\beta_1^N}$, 即对 $n=N$ 成立。

设对 n 时成立 ($0 < n \leq N$)。

$\forall \pi^{(n-1)} \in \bar{F}_{n-1} \subset e(F \times \bar{F}_n)$, 显然 $\bar{C}_{n-1}(\pi^{(n-1)}) \not\geq \frac{v^*}{\beta_1^{n-1}}$, 由归纳假设 $\pi^{(n-1)} \in F \times \bar{F}_n \subset F \times F_n$ 。

令 $\tilde{\pi}^{(n-1)} \in F \times F_n$ 且 $V_{n-1}(\pi^{(n-1)}) \leq V_{n-1}(\tilde{\pi}^{(n-1)})$ 由 (11), $\bar{C}_{n-1}(\pi^{(n-1)}) \geq \bar{C}_{n-1}(\tilde{\pi}^{(n-1)})$, 令 $\tilde{\pi}^{(n-1)} = (\tilde{f}, \tilde{\pi}^{(n)})$, 则 $\bar{C}_{n-1}(\tilde{\pi}^{(n-1)}) = [\tilde{f}c] + [AK\tilde{f} \sim \beta a q \bar{C}_n(\tilde{\pi}^{(n)})] \geq [AK\tilde{f} \sim \beta a q \bar{C}_n(\pi^{(n)})]$, 若 $\bar{C}_n(\pi^{(n)}) > \frac{v^*}{\beta_1^n}$, 则 $\bar{C}_{n-1}(\pi^{(n-1)}) \geq \bar{C}_{n-1}(\pi^{(n-1)}) \geq \frac{v^*}{\beta_1^{n-1}}$ 。这与 $\bar{C}_{n-1}(\pi^{(n-1)}) \not\geq \frac{v^*}{\beta_1^{n-1}}$ 矛盾。故 $\bar{C}_{n-1}(\pi^{(n-1)}) \not\geq \frac{v^*}{\beta_1^{n-1}}$ 由归纳

假设, $\exists \pi^{(n)} \in \bar{F}_n$, 满足 $V_n(\pi^{(n)}) \leq V_n(\pi_1^{(n)})$ 令 $\pi_1^{(n-1)} = (\tilde{f}, \pi_1^{(n)}) \in F \times \bar{F}_n$, 易见

$$V_{n-1}(\pi^{(n-1)}) \leq V_{n-1}(\tilde{\pi}^{(n-1)}) \leq V_{n-1}(\pi_1^{(n-1)})$$

由于 $\pi^{(n-1)} \in e(F \times \bar{F}n)$, 我们有 $V_{n-1}(\pi^{(n-1)}) = V_{n-1}(\tilde{\pi}^{(n-1)}) = V_{n-1}(\tilde{\pi}_1^{(n-1)})$
 所以 $\pi^{(n-1)} \in e(F \times \bar{F}n) = F_{n-1}$, 即 $\bar{F}_{n-1} \subset F_{n-1}$ 。

$\forall \pi^{(n-1)} \in \Pi(n-1)$, 令 $\bar{C}_{n-1}(\pi^{(n-1)}) \not\geq \frac{v^*}{\beta_1^*} 1$ 及 $\pi^{(n-1)} = (f, \pi^{(n)})$, 则 $\bar{C}_{n-1}(\pi^{(n-1)}) = [fc] + [f\beta aq$
 $\bar{C}_n(\pi^{(n)})] \geq [f\beta aq \bar{C}_n(\pi^{(n)})]$

若 $\bar{C}_n(\pi^{(n)}) > \frac{v^*}{\beta_1^*} 1$, 则 $\bar{C}_{n-1}(\pi^{(n-1)}) > \frac{v^*}{\beta_1^*} 1$, 得到矛盾。故有 $\bar{C}_{n-1}(\pi^{(n-1)}) \not\geq \frac{v^*}{\beta_1^*} 1$ 。由归纳假设,
 $\exists \tilde{\pi}^{(n)} \in \bar{F}n$, 满足 $V_n(\pi^{(n)}) \leq V_n(\pi_1^{(n)})$, 令 $\tilde{\pi}^{(n-1)} = (f, \tilde{\pi}^{(n)}) \in F \times \bar{F}n$, 易知 $V_{n-1}(\pi^{(n-1)}) \leq V_{n-1}$
 $(\tilde{\pi}^{(n-1)})$, 由引理 3.1, $\exists \pi_2^{(n-1)} \in e(F \times \bar{F}n)$, 满足 $V_{n-1}(\pi^{(n-1)}) \leq V_{n-1}(\tilde{\pi}_1^{(n-1)})$ 。故 $V_{n-1}(\pi^{(n-1)}) \leq V_{n-1}$
 $(\pi_1^{(n-1)})$, 由(11)我们得到 $\bar{C}_{n-1}(\tilde{\pi}_1^{(n-1)}) \leq \bar{C}_{n-1}(\pi^{(n-1)})$ 。因 $\bar{C}_{n-1}(\pi_1^{(n-1)}) \not\geq \frac{v^*}{\beta_1^*} 1$ 有 $\bar{C}_{n-1}(\tilde{\pi}_1^{(n-1)}) \not\geq \frac{v^*}{\beta_1^*} 1$, 即
 $\tilde{\pi}_1^{(n-1)} \in \bar{F}_{n-1}$ 。

令 $\pi \in \Pi(0)$, $\bar{C}_0(\pi) \leq v^*$, 又令 $\pi = (f_0, f_1, \dots, f_{n-2}, \pi^{(n-1)})$, 类似上述证明, 有 $\bar{C}_{n-1}(\pi_1^{(n-1)}) \not\geq$
 $\frac{v^*}{\beta_1^*} 1$ 。由上面证明, $\exists \tilde{\pi}_1^{(n-1)} \in \bar{F}(n-1)$, 满足 $V_{n-1}(\pi^{(n-1)}) \leq V_{n-1}(\tilde{\pi}_1^{(n-1)})$ 即 $\bar{F}_{(n-1)} \neq \emptyset$ 。

结合起来, 已证对 $n-1$ 时成立。#。

令 $\bar{W}_1 = \{\pi \in \bar{F}_0 \mid \pi \text{ 为 } FCSMDP \text{ 的最优策略 (对 } \bar{F}_0)\}$ 。

$\epsilon > 0$, 令 $\bar{W}_2 = \{\pi \in \bar{F}_0 \mid \pi \text{ 为 } FCSMDP \text{ 的 } \epsilon\text{-最优策略 (对 } \bar{F}_0)\}$ 。

定理 3.4

(i) 若 $\pi \in \bar{W}_1$, 则 π 为 $FCSMDP$ 的最优策略 (对 $\Pi(0)$); 若 $\bar{W}_1 = \emptyset$, $FCSMDP$ 没有最优策
 略 (对 $\Pi(0)$)。

(ii) 若 $\pi \in \bar{W}_2$, 则 π 为 $FCSMDP$ 的 ϵ 最优策略 (对 $\Pi(0)$); 若 $\bar{W}_2 = \emptyset$, $FCSMDP$ 没有 ϵ -最
 优策略 (对 $\Pi(0)$)。

(iii) 若 $\{\pi \in \bar{F}_0 \mid \bar{C}_0(\pi) \leq v\} = \emptyset$, 则 $U = 0$ 。

证明 类似于定理 3.2 的证明

4 一个算法

对 $\epsilon, \epsilon' > 0$, 取正整数 N 满足

$$M_1 \beta^{(N+1)} \frac{1}{1-\beta} \leq \frac{\epsilon}{4} \quad M_2 \beta^{(N+1)} \frac{1}{1-\beta} \leq \frac{\epsilon'}{2}$$

其中, β 定义见(4)式。由约束 $SMDP$ 和 N , 我们可得到一个 $FCSMDP: \{S, A, q, t, r, c, \bar{v}, N\}$,
 其中, $\bar{v} = v + \frac{\epsilon'}{2} 1$, 其它均同前。可用 3 中方法求解本模型。

定理 4.1 令 $\pi^{(0)} = (f_0, f_1, \dots, f_N) \in \Pi(0)$ 为 $FCSMDP$ 的 $\frac{\epsilon}{2}$ -最优策略 (对 $\Pi(0)$), $\forall g_i$
 $\in F, i = N+1, N+2, \dots$, 则 $\pi^* = (f_0, f_1, \dots, f_N, g_{(N+1)}, g_{(N+2)}, \dots)$ 为带约束 $SMDP$ 的 (ϵ, ϵ') -
 最优策略。

证明 令 $(g_{(N+1)}, \dots) = \pi^{(N+1)}$, 易见

$$\bar{C}_0(\pi^{(0)}) \leq \bar{v} = v + \frac{\epsilon}{2} 1$$

$$\text{则 } C(\pi^*) = \int_0^\infty e^{-\alpha T} C(dT \mid \pi^*) = \sum_{k=0}^\infty [\pi^* \beta_k q c]_k$$