第 24 卷第 4 期　　南京师大学报(自然科学版)　　Vol.24 No.4
2001 年　　JOURNAL OF NANJING NORMAL UNIVERSITY(Natural Science)　　2001

有向图上的动态规划

许扬灵

(中国人民解放军理工大学理学院,南京,211101)

[摘要]　在集合论的基础上将离散的动态规划形式化,用递归函数刻划了动态规划的目标函数,并在有向图上建立了动态规划.推广了动态规划方法的应用范围.

[关键词]　有向图 ;动态规划 ;目标函数 ;递归函数

[中图分类号]O221.3 ;　[文献标识码]A ;　[文章编号]1001-4616(2001)04-0020-04

0　引言

动态规划是由 Richard Bellman 于 1957 年提出的.它是求解多阶段决策问题的有效方法.然而,正如文献[1]中指出的那样;"从严格的数学意义上来说,动态规划还不是一个很完善的数学分支".至今,动态规划的理论仍建立在以下这些基本概念之上 :级、状态、决策、策略和收益等.很明显,这些概念还不是严格意义上的数学概念.文[1]给出了最优性定理的证明.这个定理只是对目标函数为相加型的可分函数的动态规划而言的.实际上,动态规划的目标函数可以是多种多样的.许多人研究了可用 Bellman 原理求解的动态规划的目标函数的特征,但结果不甚完善.本文用递归函数刻划了动态规划的目标函数,并在简单有向图上建立了动态规划.因此,推广了动态规划的应用范围.

1　多阶段有向图上的动态规划

定义 1　设 $D=(V,E)$ 是一个简单有向图,如果存在两两不相交的非空集合 A_i,$i=0,1,\cdots,N$,使得

$$V=\bigcup_{i=0}^{N}A_i,$$

并且对于任意有向边 $e=(u,v)\in E$,存在自然数 k,使得

$$u\in A_{k-1},v\in A_k,$$

对于任意顶点 $u\in V$,u 都在某条从 x_0 到 x_N 的通路上,其中 $x_0\in A_0$,$x_N\in A_N$,那么,我们称这个简单有向图 $D=(V,E)$ 为一个多阶段有向图.

我们假设 A_0 为一个单点集 $\{x_0\}$,A_N 也为一个单点集 $\{x_N\}$.本文中符号 x_k 表示它是 A_k 中的某个顶点.$W[x_0,x_k]$ 表示所有从顶点 x_0 到顶点 x_k 的通路的集合.根据上面的定义,这种通路必为路径.设 f 是一个从 $w[x_0,x_N]$ 到实数集 $\mathbf{R}$ 中的映射.则我们有最优化问题:

$$\min f(w)$$

收稿日期:2000-12-22
作者简介:许扬灵,1959—,中国人民解放军理工大学理学院讲师,硕士,主要从事应用数学的教学与研究.

$s.t. \quad w \in W[x_0, x_N],$ (DP$^-$)

或者 $\max f(w)$

$s.t. \quad w \in W[x_0, x_N].$ (DP$^+$)

为了使用动态规划原理求解以上问题,我们要把 $f(w)$ 的定义域扩大. 把原始问题嵌入到一类相似的问题中去. 令

$$\Gamma = \bigcup_{x_k \in V} W[x_0, x_k].$$

Γ 中的元素用 $(x_0, x_1, x_2, \cdots, x_k)$ 表示. Γ 中有一条长度为 0 的路径,用 (x_0) 表示. 这是下文中递归函数所需要的.

把 $f(w), w \in W[x_0, x_N]$ 的定义域扩大到 Γ,这在某些实际问题中是十分自然的. 如求赋权图上给定两点之间的最短路径,而最短路径也可用于其它两点. 而在另外一些问题上,这种推广则需要一定的技巧.

设 $\varphi(s, x_{k-1}, x_k, k)$ 是一个实值函数,其中,$s \in \mathbf{R}$,k 是一个自然数,$1 \leqslant k \leqslant N$,$x_{k-1} \in A_{k-1}$,$x_k \in A_k$. 为了使目标函数可以应用动态规划原理,我们还要求函数 φ 满足单调性要求:当 k 取定时,$\varphi(s, x_{k-1}, x_k, k)$ 作为 s 的函数,不管 x_{k-1}, x_k 取何值,它总具有一种单调性. 当它是单调增加时,令 $\lambda(k) = +1$,当它是单调减少时,令 $\lambda(k) = -1$.

定义 2 如果对于任意路径

$$(x_0, x_1, \cdots, x_k) \in \Gamma,$$

函数 $f(x_0, x_1, \cdots, x_k)$ 由下式递归地给出:

$$f(x_0, x_1, \cdots, x_k) = \varphi(f(x_0, x_1, \cdots, x_{k-1}), x_{k-1}, x_k, k),$$

$f(x_0)$ 为一已知实数. 那么,我们称 $f(w)$ 为关于有向图 (D, x_0, x_N) 的一个 Bellman 函数. $\varphi(s, x_{k-1}, x_k, k)$ 称为 $f(w)$ 的发生函数. 当 $f(w)$ 为一个 Bellman 函数时,我们称 (DP$^-$) 或 (DP$^+$) 为一个动态规划问题.

一个动态规划问题可以记为 (D, x_0, x_N, f). 其中 D 称为基础图. x_0, x_N 分别称为始点和终点,$f(n)$ 称为目标函数,N 称为阶段数,A_k 称为 k 级状态集,V 称为状态空间,$W[x_0, x_N]$ 称为策略空间.

我们规定 $opt(+1)$ 表示 max,$opt(-1)$ 表示 min.

定理 对于定义 2 提出的动态规划问题,它的最优性递归方程为

$$\underset{(x_0,\cdots,x_k)\in W[x_0,x_k]}{opt(\alpha_k)} f(x_0, \cdots, x_k) = \underset{X_{R-1}\in A_{R-1}}{opt}(\alpha_k) \{\varphi(\underset{(x_0,\cdots,x_{k-1})\in W[x_0,x_{k-1}]}{opt(\alpha_k s(k))} f(x_0, \cdots, x_{k-1}), k_{k-1, x_k}, k)\}.$$

其中

$$\alpha_k = \begin{cases} r\lambda(N)\lambda(N-1)\cdots\lambda(k+1), & k < N \\ r, & k = N \end{cases}$$

证明 对于 (DP) 问题

$opt(r) \quad f(w)$

$s.t. \quad w \in T,$

根据 $f(w)$ 的递归结构和 $\varphi(s, x_{k-1}, x_k, k)$ 的单调性,我们有

$$\underset{(x_0,\cdots,x_N)\in W[x_0,x_N]}{opt(r)} f(x_0, \cdots, x_N)$$

$$= \underset{(x_0,\cdots,x_N)\in W[x_0,x_N]}{opt(r)} f(x_0, \cdots, x_N)$$

$$= \mathop{opt(r)}_{(x_0,\cdots,x_N)\in W[x_0,x_N]} \varphi(f(x_0,\cdots,x_{N-1}),x_{N-1},x_N,N)$$
$$= \mathop{opt}_{x_{N-1}\in A_{N-1}} \{\mathop{opt(r)}_{(x_0,\cdots,x_{N-1})\in W(x_0,x_{N-1})} \varphi(f(x_0,\cdots,x_{N-1}),x_{N-1},x_N,N)\}$$
$$= \mathop{opt}_{x_{N-1}\in A_{N-1}} \{\varphi \mathop{opt(r\lambda(N))}_{(x_0,\cdots,x_{N-1})\in W(x_0,x_{N-1})} f(x_0,\cdots,x_{N-1}),x_{N-1},x_N,N)\}$$

上式中的第二个优化符要寻找的路径的长为 $N-1$,称之为第$(N-1)$个路径优化符.

下面略去部分优化符下面的约束.

$$opt(r\lambda(N))f(x_0,\cdots,x_{N-1})$$
$$= opt(r\lambda(N))\varphi(f(x_0,\cdots,x_{N-2}),x_{N-2},x_{N-1},N-1)$$
$$= \mathop{opt(r\lambda(N))}_{x_{N-2}\in A_{N-2}} \{opt(r\lambda(N))\varphi(f(x_0,\cdots,x_{N-2}),x_{N-2},x_{N-1},N-1)\}$$
$$= \mathop{opt(r\lambda(N))}_{x_{N-2}\in A_{N-2}} \{\varphi(opt(r\lambda(N)\lambda(N-1)f(x_0,\cdots,x_{N-2}),x_{N-1},N-1)\}.$$

上式中出现的路径优化符是第$(N-2)$个.

应用数学归纳法,可以得到第 $k(1\leqslant k\leqslant N)$ 个路径优化符为 $opt(\alpha_k)$,其中

$$\alpha_k = \begin{cases} r\lambda(N)\lambda(N-1)\cdots\lambda(k+1), & k<N, \\ r, & k=N. \end{cases}$$

一般地,我们有

$$\mathop{opt(\alpha_k)}_{(x_0,\cdots,x_k)\in W[x_0,x_k]} f(x_0,\cdots,x_k)$$
$$= \mathop{opt(\alpha_k)}_{x_{k-1}\in A_{k-1}} \{\varphi(\mathop{opt(\alpha_k\lambda(k))}_{(x_0,\cdots,x_{k-1})\in W[x_0,x_{k-1}]} f(x_0,\cdots,x_{k-1}),x_{k-1},x_k,k)\}.$$

2　例子和算法

下面是三个 Bellman 函数的例子.其中 x,y 为赋权顶点所对应的数值.

例 1　令 $\varphi(s,x,y,k)=S+\psi_k(y)$,$\psi_k(y)$ 为已知函数,$f(x_0)=0$,则

$$f(x_0,x_1,\cdots,x_k)=\sum_{i=1}^{k}\psi_i(x_i).$$

例 2　令 $\varphi(s,x,y,k)=s+xy$,$f(x_0)=0$,则

$$f(x_0,x_1,\cdots,x_k)=x_0x_1+x_1x_2+\cdots+x_{k-1}x_k.$$

例 3　令 $\varphi(s,x,y,k)=s\cdot\sin\frac{k\pi}{3}+y-x$,$f(x_0)=1$,则

$$f(x_0,x_1)=x_1-x_0+\frac{\sqrt{3}}{2};$$

$$f(x_0,x_1,x_2)=x_2-(1-\frac{\sqrt{3}}{2})x_1-\frac{\sqrt{3}}{2}x_0+\frac{3}{4}.$$

当动态规划问题中的基础图 D、发生函数 $\varphi(s,x,y,k)$ 以及 $f(x_0)$ 确定后,可用下面的算法求解这个问题.

多阶段有向图上的动态规划的算法步骤:

(1) 对 $k=1,2,\cdots,N$,计算

$$\alpha_k = \begin{cases} rS(N)S(N-1)\cdots S(k+1), & k<N \\ r, & k=N \end{cases}$$

令 $k:=1$.

(2) 对任一 $x_k \in A_k$，求解

$$\underset{(x_0,\cdots,x_k)\in W[x_0,x_k]}{opt(\alpha_k)} f(x_0,\cdots,x_k)$$

$$= \underset{x_{k-1}\in A_{k-1}}{opt(\alpha_k)} \{ \underset{(x_0,\cdots,x_{k-1})\in W[x_0,x_{k-1}]}{opt(\alpha_{k-1})} \varphi(f(x_0,\cdots,x_{k-1}),x_{k-1},x_k,k)\},$$

并保存所求得的最优值及其路径.

(3) 检查 k 是否为 N，若是，算法终止. $\underset{x_{N-1}\in A_{N-1}}{opt(\alpha_N)} f(x_0,\cdots,x_N)$ 即为原问题的最优值；否则，令 $k := k+1$，转(2).

当 $f(n)$ 取得(DP)的最优值时，w 称为(DP)的最优路径(即最优策略). 这条最优路径可以通过逆序跟踪求得.

[参考文献]

[1] 张有为. 动态规划[M]. 长沙：湖南科学出版社，1991.

[2] 耿素云，方新贵. 离散数学[M]. 北京：北京大学出版社，1989.

[3] Bellman R E. Dynamic Programming[M]. Princeton：Princeton University Press，1957.

Dynamic Programming on a Directed Graph

Xu Yangling

(PLA University of Science and Technology，Nanjing，211101，PRC)

Abstract：Dynamic programming is built up on a basis of set theory. Its criterion functions are described by recursive functions. We build up dynamic programming on a directed graph. So the range of applications of dynamic programming should be widened.

Key words：divected graph；dynamic programming；objective function；recursive function

[责任编辑：陆炳新]