现代管理科学方法
（第4讲）

郭仁拥 博士/教授/博导

讲授内容

1. 考虑路段流量约束的交通分配问题
2. 求解问题的增广拉格朗日对偶算法
3. 算法实施和迭代过程
4. 数值算例

1. 考虑路段流量约束的交通分配问题

为什么要在交通分配模型中考虑路段流量约束？

现实道路路段只可以承载有限交通流量，为了使模型能够更精确描述和预测现实交通状况，引入路段流量约束

考虑路段流量约束的交通分配问题可以刻画为如下优化模型

$$\min_{(\mathbf{x},\mathbf{f})} V(\mathbf{x}) = \sum_{a\in L} \int_0^{x_a} c_a(x)\mathrm{d}x \quad (1)$$

$$s.t.,\quad x_a = \sum_{w\in W} \sum_{r\in R_w} \delta_{a,r} f_{r,w},\quad \forall\ a\in L \quad (2)$$

$$d_w = \sum_{r\in R_w} f_{r,w},\quad \forall\ w\in W \quad (3)$$

$$f_{r,w} \ge 0,\quad \forall\ r\in R_w,\ w\in W \quad (4)$$

$$x_a \le u_a,\quad \forall\ a\in L \quad (5)$$

考虑路段流量约束的交通分配问题可以刻画为如下优化模型

$$\min_{(\mathbf{x},\mathbf{f})} V(\mathbf{x}) = \sum_{a\in L}\int_0^{x_a} c_a(x)\mathrm{d}x \qquad (1)$$

$$s.t., \quad x_a = \sum_{w\in W}\sum_{r\in R_w}\delta_{a,r} f_{r,w}, \quad \forall\, a\in L \qquad (2)$$

$$d_w = \sum_{r\in R_w} f_{r,w}, \quad \forall\, w\in W \qquad (3)$$

$$f_{r,w} \geq 0, \quad \forall\, r\in R_w,\ w\in W \qquad (4)$$

传统交通分配问题

$$x_a \leq u_a, \quad \forall\, a\in L \qquad (5)$$

考虑路段流量约束的交通分配问题可以刻画为如下优化模型

$$\min_{(\mathbf{x},\mathbf{f})} V(\mathbf{x}) = \sum_{a\in L}\int_0^{x_a} c_a(x)\mathrm{d}x \qquad (1)$$

$$s.t., \quad x_a = \sum_{w\in W}\sum_{r\in R_w}\delta_{a,r} f_{r,w}, \quad \forall\, a\in L \qquad (2)$$

$$d_w = \sum_{r\in R_w} f_{r,w}, \quad \forall\, w\in W \qquad (3)$$

$$f_{r,w} \ge 0, \quad \forall\, r\in R_w,\ w\in W \qquad (4)$$

$$\boxed{x_a \le u_a, \quad \forall\, a\in L} \qquad (5)$$

路段流量约束条件

u_a为路段a的通行能力上限

设 $\mathbf{u} = (u_a, a \in L)^{\mathrm{T}}$，模型又可以简写为

$$\min_{(\mathbf{x},\mathbf{f})} V(\mathbf{x}) = \sum_{a \in L} \int_0^{x_a} c_a(x)\mathrm{d}x \tag{6}$$

$$s.t., \quad (\mathbf{x}, \mathbf{f}) \in \Omega \equiv \left\{(\mathbf{x}, \mathbf{f}) \mid \mathbf{x} = \mathbf{\Delta f}, \mathbf{d} = \mathbf{\Lambda f}, \mathbf{f} \geq \mathbf{0}\right\} \tag{7}$$

$$\mathbf{x} \leq \mathbf{u} \tag{8}$$

假设优化模型的可行集是非空的，即考虑路段流量约束的网络可以承载下所有的出行需求

目标函数关于 **x** 是严格凸的（基于假设路段费用函数是增函数），且可行集是非空闭凸集，因此优化模型(1)~(5)（或模型(6)~(8)）是一个严格凸优化问题

因此，该优化问题的最优路段流量解是唯一的

$(\mathbf{x}^*, \mathbf{f}^*)$是优化模型(1)~(5)的一个最优解当且仅当它满足如下KKT条件：

$$\frac{\partial V(\mathbf{x})}{\partial f_{r,w}} = C_{r,w}(\mathbf{f}) = \mu_w + \lambda_{r,w} - \sum_{a \in L} \beta_a \delta_{a,r}\ ,\quad \forall\ r \in R_w,\ w \in W$$

$$\lambda_{r,w} \geq 0\,,\quad f_{r,w} \geq 0\,,\quad \lambda_{r,w} f_{r,w} = 0\,,\quad \forall\ r \in R_w,\ w \in W$$

$$\beta_a \geq 0\,,\quad u_a - x_a \geq 0\,,\quad \beta_a (u_a - x_a) = 0\,,\quad \forall\ a \in L$$

其中 $\boldsymbol{\mu} = (\mu_w, w \in W)^{\mathrm{T}}$ 、 $\boldsymbol{\lambda} = (\lambda_{r,w}, r \in R_w, w \in W)^{\mathrm{T}}$ 和 $\boldsymbol{\beta} = (\beta_a, a \in L)^{\mathrm{T}}$ 分别是约束条件(3)、(4)和(5)的拉格朗日乘子。

令广义路径出行费用

$$\bar{C}_{r,w} = C_{r,w}(\mathbf{f}) + \sum_{a \in L} \beta_a \delta_{a,r} = \sum_{a \in L} \left(c_a(x_a) + \beta_a \right) \delta_{a,r}$$

以上的KKT条件也可以写作

$$\bar{C}_{r,w} - \mu_w \geq 0, \quad f_{r,w} \geq 0, \quad \forall\ r \in R_w,\ w \in W$$

$$\left(\bar{C}_{r,w} - \mu_w \right) f_{r,w} = 0, \quad \forall\ r \in R_w,\ w \in W$$

$$\beta_a \geq 0, \quad u_a - x_a \geq 0, \quad \beta_a (u_a - x_a) = 0, \quad \forall\ a \in L$$

以上的KKT条件如何解释？提示：$\boldsymbol{\beta} = (\beta_a, a \in L)^{\mathrm{T}}$ 可解释为排队延误或收费

增加路段流量约束后，问题并没有变的多么复杂，但问题的求解难度增加了，不能直接使用Frank-Wolfe（FW）算法求解，为什么？

增加了路段流量约束后，在每次算法迭代中就不能用全有全无流量分配方法来计算目标函数下降方向了

求解思路：由于无路段流量约束的交通分配问题可以被有效求解，因此可以将有路段流量约束的模型转换为一系列无路段流量约束的子问题来求解，使得无路段流量约束问题逼近于有路段流量约束问题。子问题可以由FW算法来求解。

2. 求解问题的增广拉格朗日对偶算法

如果使用外惩罚法（Exterior penalty method），将路段流约束 $\mathbf{x} \le \mathbf{u}$ 放到目标函数中，惩罚函数 $P: \mathrm{R}^{|L|} \to \mathrm{R}$ 满足

(1) $P(\mathbf{x}) \ge 0$， $\forall\ \mathbf{x} \in \Omega_{\mathrm{L}} \equiv \left\{ \mathbf{x} \mid (\mathbf{x}, \mathbf{f}) \in \Omega \right\}$

(2) $P(\mathbf{x}) = 0$ 当且仅当 $\mathbf{x} \le \mathbf{u}$

(3) 函数 P 在集合 Ω_{L} 上是连续的

一个例子是 $$P(\mathbf{x}) = \sum_{a \in L} r_a [x_a - u_a]_+^{m_a}, \tag{9}$$

这里 $[\cdot]_+ = \max\{\cdot, 0\}$， $r_a > 0$ 和 $m_a \ge 2$

引入惩罚参数$\gamma > 0$，惩罚目标函数可写为

$$P_\gamma(\mathbf{x}) = V(\mathbf{x}) + \gamma P(\mathbf{x}) \tag{10}$$

惩罚子问题写作 $P_\gamma = \min_{\mathbf{x}\in\Omega_\mathrm{L}} P_\gamma(\mathbf{x})$，该子问题的解表示为

$\mathbf{x}(\gamma) = \arg\min_{\mathbf{x}\in\Omega_\mathrm{L}} P_\gamma(\mathbf{x})$，它们满足

$$P_\gamma \le V(\mathbf{x}^*)\text{，} \forall\ \gamma > 0\text{；}\ \lim_{\gamma\to+\infty} \mathbf{x}(\gamma) = \mathbf{x}^*$$

这里$\mathbf{x}^*$是优化问题(1)~(5)的唯一最优解

惩罚约束的最优拉格朗日乘子$\boldsymbol{\beta}^*$可由如下公式来估计

$$\lim_{\gamma \to +\infty} \gamma \nabla P(\mathbf{x})\Big|_{\mathbf{x}=\mathbf{x}(\gamma)} = \boldsymbol{\beta}^*$$

为了避免罚函数法的病态问题，可以在惩罚目标函数(10)中引入一个拉格朗日乘子项，这样就产生一个增广拉格朗日函数

惩罚函数(9)中的$r_a = 1/2$和$m_a = 2$，$\forall\ a \in L$，增广拉格朗日函数成为

$$L_\gamma(\mathbf{x},\boldsymbol{\beta}) = V(\mathbf{x}) + \frac{1}{2\gamma}\sum_{a \in L}\left(\left[\beta_a + \gamma(x_a - u_a)\right]_+^2 - \beta_a^2\right) \qquad (11)$$

无路段流量约束的交通分配子问题表示为

$$L_\gamma(\boldsymbol{\beta}) = \min_{\mathbf{x}\in\Omega_\mathrm{L}} L_\gamma(\mathbf{x},\boldsymbol{\beta}) \tag{12}$$

子问题的解表示为

$$\mathbf{x}(\boldsymbol{\beta},\gamma) = \arg\min_{\mathbf{x}\in\Omega_\mathrm{L}} L_\gamma(\mathbf{x},\boldsymbol{\beta}) \tag{13}$$

它们满足(1) $L_\gamma(\boldsymbol{\beta}) \le L_\gamma(\boldsymbol{\beta}^*) = V(\mathbf{x}^*)$，$\forall\ \boldsymbol{\beta} \ge \mathbf{0}$

(2) $\lim_{\boldsymbol{\beta}\to\boldsymbol{\beta}^*} \mathbf{x}(\boldsymbol{\beta},\gamma) = \mathbf{x}(\boldsymbol{\beta}^*,\gamma) = \mathbf{x}^*$

这里 $\mathbf{x}^*$ 是优化问题(1)~(5)的唯一最优解，$\boldsymbol{\beta}^*$ 是约束(5)的最优拉格朗日乘子

使用增广拉格朗日罚函数的优点：对于任意的$\gamma \geq 0$，由任意的拉格朗日乘子最优值，增广拉格朗日对偶目标函数能被最大化；因此，当惩罚参数取有限值，优化模型(1)~(5)的最优解可以被计算得到

通过求解如下增广拉格朗日对偶问题，可以得到最优乘子

$$\max_{\boldsymbol{\beta} \geq \mathbf{0}} L_\gamma(\boldsymbol{\beta}) \tag{14}$$

这里函数L_γ是凹的和可微的，它的偏导数为

$$\frac{\partial L_\gamma(\boldsymbol{\beta})}{\partial \beta_a} = \max\left\{\left(x_a(\boldsymbol{\beta}, \gamma) - u_a\right), -\frac{\beta_a}{\gamma}\right\}, \quad \forall\ a \in L$$

为求解对偶问题(14)，拉格朗日乘子$\boldsymbol{\beta}$可以沿着正梯度方向以步长γ更新，即

$$\beta_a \equiv \beta_a + \gamma \frac{\partial L_\gamma(\boldsymbol{\beta})}{\partial \beta_a} = \left[\beta_a + \gamma\left(x_a(\boldsymbol{\beta},\gamma) - u_a\right)\right]_+, \quad \forall\ a \in L$$

步长γ取值太小会导致算法收敛速度太慢；取值太大会导致拉格朗日子问题的病态问题

步长γ可按照如下策略进行更新：选取较小的γ初始值；当对偶约束的不可行程度没有足够快的改进，则γ值增加

因此，引入惩罚参数的非减序列$\{\gamma^k\}$，路段流量和拉格朗日乘子按如下公式进行迭代

$$\mathbf{x}^k = \mathbf{x}(\boldsymbol{\beta}^k, \gamma^k) \tag{15}$$

$$\beta_a^{k+1} = \left[\beta_a^k + \gamma^k \left(x_a^k - u_a\right)\right]_+, \quad \forall\, a \in L \tag{16}$$

对于$k = 1, 2, \cdots$

3. 算法实施和迭代过程

初始拉格朗日乘子选取：初始乘子选取对算法计算效率有显著的影响

设 $\mathbf{x}^0$ 是无路段流量约束交通分配问题的均衡解，设置

$$\beta_a^1 = \begin{cases} c_a(x_a^0) - c_a(u_a), & \text{如果} a \in \overline{L}^0 \\ 0, & \text{如果} a \notin \overline{L}^0 \end{cases} \tag{17}$$

这里 $\overline{L}^0 = \{a \in L \mid x_a^0 > u_a\}$

以上是启发式设置方法，针对流量超过通行能力上限的路段，施加一个额外的成本

惩罚参数的选取：取值太小导致算法收敛速度太慢；取值太大导致拉格朗日子问题的病态问题

原则：算法迭代中，如果对偶约束的不可行程度没有足够的下降，则惩罚参数值增加

设 $g_a(x_a,\beta_a,\gamma)=\max\left\{(x_a-u_a),-\beta_a/\gamma\right\}$，惩罚参数值按如下规则进行更新

$$\gamma^{k+1}=\begin{cases}\kappa\gamma^k, & \text{如果}\left\|\mathbf{g}(\mathbf{x}^k,\boldsymbol{\beta}^k,\gamma^k)\right\|>\eta\left\|\mathbf{g}(\mathbf{x}^{k-1},\boldsymbol{\beta}^{k-1},\gamma^{k-1})\right\| \\ \gamma^k, & \text{否则}\end{cases} \tag{18}$$

这里 $\mathbf{g}(\mathbf{x},\boldsymbol{\beta},\gamma)=\left(g_a(x_a,\beta_a,\gamma),a\in L\right)^{\mathrm{T}}$，建议取值 $2\le\kappa\le10$ 和 $\eta=0.25$

子问题的求解：在每次算法迭代中，可利用FW算法来求解无路段流量约束的交通分配子问题(12)，得到流量 $\mathbf{x}^k$

算法停止准则：算法终止准则是

$$\|\boldsymbol{\beta}^{k+1}-\boldsymbol{\beta}^{k}\|<\varepsilon \tag{19}$$

其中 $\varepsilon(>0)$是一个提前确定的参数

接下来证明条件(19)能被作为算法的一个停止准则，即当

$\boldsymbol{\beta}^{k+1}=\boldsymbol{\beta}^{k}$ 时，$(\mathbf{x}^{k},\mathbf{f}^{k})$ 是优化问题(1)~(5)的一个最优解

首先，优化问题(12)是一个严格凸优化问题，因此该问题的KKT条件是充要条件

$(\mathbf{x}^k, \mathbf{f}^k)$ 是优化问题(12)的一个最优解当且仅当 $(\mathbf{x}^k, \mathbf{f}^k)$ 满足如下KKT条件

$$\frac{\partial L_\gamma(\mathbf{x}, \boldsymbol{\beta})}{\partial f_{r,w}} = \sum_{a\in L} \frac{\partial V(\mathbf{x})}{\partial x_a} \frac{\partial x_a}{\partial f_{r,w}} + \sum_{a\in L} \left[\beta_a + \gamma(x_a - u_a)\right]_+ \frac{\partial x_a}{\partial f_{r,w}}$$

$$= \sum_{a\in L} \left(c_a(x_a) + \left[\beta_a + \gamma(x_a - u_a)\right]_+\right) \delta_{a,r}$$

$$= \mu_w + \lambda_{r,w}, \quad \forall\ r \in R_w,\ w \in W \tag{20}$$

$$\lambda_{r,w} \geq 0, \quad f_{r,w} \geq 0, \quad \lambda_{r,w} f_{r,w} = 0, \quad \forall\ r \in R_w,\ w \in W \tag{21}$$

其次，当$\boldsymbol{\beta}^{k+1}=\boldsymbol{\beta}^{k}$时，公式(16)成为

$$\beta_a^k=\left[\beta_a^k+\gamma^k\left(x_a^k-u_a\right)\right]_+,\ \forall\ a\in L \tag{22}$$

也就是说，当$\beta_a^k=0$时，$x_a^k\le u_a$；当$\beta_a^k>0$时，$x_a^k=u_a$

因此有，

$$\beta_a^k\ge 0,\ u_a-x_a^k\ge 0,\ \beta_a^k(u_a-x_a^k)=0,\ \forall\ a\in L \tag{23}$$

再次，由子问题(15)和公式(20)~(22)可知，$(\mathbf{x}^k, \mathbf{f}^k)$ 满足

$$\sum_{a\in L}\left(c_a(x_a)+\beta_a\right)\delta_{a,r}-\mu_{r,w}\geq 0,\quad f_{r,w}\geq 0,\quad \forall\, r\in R_w,\ w\in W \tag{24}$$

$$\left(\sum_{a\in L}\left(c_a(x_a)+\beta_a\right)\delta_{a,r}-\mu_{r,w}\right)f_{r,w}=0,\quad \forall\, r\in R_w,\ w\in W \tag{25}$$

由KKT条件，公式(23)~(25)表明$(\mathbf{x}^k, \mathbf{f}^k)$是优化问题(1)~(5)的一个最优解

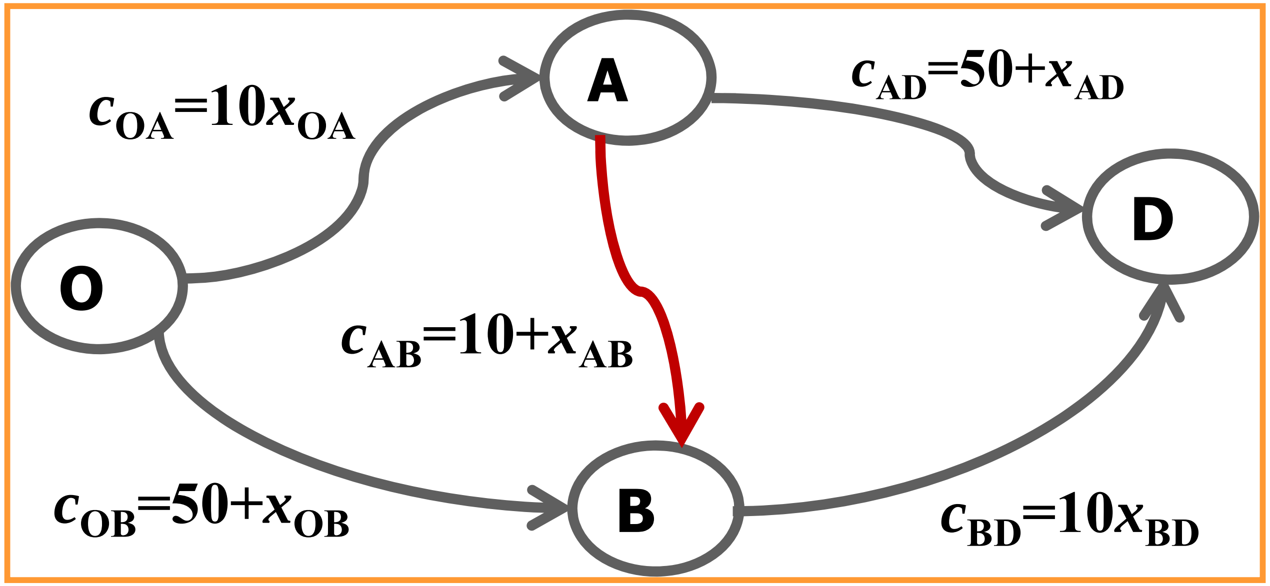
算法迭代过程：第1步（初始化）：利用FW算法求解无流量约束交通分配问题，得到流量 $\mathbf{x}^0$；根据公式(17)设置 $\boldsymbol{\beta}^1$ 的值；设置 γ^1 和 ε 的值；令 $k=1$

第2步（子问题求解）：利用FW算法求解增广拉格朗日子问题(15)，得到流量 $\mathbf{x}^k$

第3步（乘子和惩罚参数更新）：由公式(16)计算乘子 $\boldsymbol{\beta}^{k+1}$，由公式(18)计算惩罚参数 γ^{k+1}

第4步（收敛性判定）：如果 $\|\boldsymbol{\beta}^{k+1}-\boldsymbol{\beta}^k\|<\varepsilon$，则算法终止；否则，设置 $k=k+1$，转到第2步

4. 数值算例



利用增广拉格朗日对偶算法求解以上网络的均衡解，其中 $d_{OD}=6$，路段流量约束为

（1） $x_{AB}\le 1$；

（2） $x_{OA}\le 3$， $x_{AD}\le 1.5$

参数取值： $\kappa=5$， $\eta=0.25$， $\gamma^1=0.1$

算例(1)

k	x_{OA}	x_{OB}	x_{AB}	x_{AD}	x_{BD}	β_{OA}	β_{OB}	β_{AB}	β_{AD}	β_{BD}	γ
0	4.00	2.00	2.00	2.00	4.00						
1	3.92	2.08	1.83	2.08	3.92	0	0	1.00	0	0	0.10
2	3.91	2.09	1.82	2.09	3.91	0	0	1.08	0	0	0.10
3	3.88	2.12	1.76	2.12	3.88	0	0	1.17	0	0	0.50
4	3.78	2.22	1.55	2.22	3.78	0	0	1.55	0	0	2.50
5	3.59	2.41	1.19	2.41	3.59	0	0	2.92	0	0	12.50
6	3.51	2.49	1.02	2.49	3.51	0	0	5.28	0	0	62.50
7	3.50	2.50	1.00	2.50	3.50	0	0	6.38	0	0	62.50
8	3.50	2.50	1.00	2.50	3.50	0	0	6.49	0	0	62.50
9	3.50	2.50	1.00	2.50	3.50	0	0	6.50	0	0	62.50
10	3.50	2.50	1.00	2.50	3.50	0	0	6.50	0	0	62.50

算例(2)

k	x_{OA}	x_{OB}	x_{AB}	x_{AD}	x_{BD}	β_{OA}	β_{OB}	β_{AB}	β_{AD}	β_{BD}	γ
0	4.00	2.00	2.00	2.00	4.00						
1	3.16	2.84	1.27	1.88	4.12	10.00	0	0	0.50	0	0.10
2	3.15	2.85	1.27	1.88	4.12	10.02	0	0	0.54	0	0.10
3	3.15	2.85	1.28	1.87	4.13	10.03	0	0	0.58	0	0.50
4	3.12	2.88	1.32	1.80	4.20	10.10	0	0	0.76	0	2.50
5	3.05	2.95	1.41	1.64	4.36	10.40	0	0	1.51	0	12.50
6	3.01	2.99	1.48	1.52	4.48	11.03	0	0	3.32	0	62.50
7	3.00	3.00	1.50	1.50	4.50	11.44	0	0	4.73	0	62.50
8	3.00	3.00	1.50	1.50	4.50	11.49	0	0	4.96	0	62.50
9	3.00	3.00	1.50	1.50	4.50	11.50	0	0	4.99	0	62.50
10	3.00	3.00	1.50	1.50	4.50	11.50	0	0	5.00	0	62.50