对有凹目标函数的一类优化问题在交通运输领域的应用

——以可靠最短路问题为例

引言：

凸优化(Convex Optimization)理论在运筹优化领域具有广泛应用。凸优化问题的局部最优解即为全局最优解，只需简单的梯度算法即可得到原问题的全局最优解。而现实生活中许多问题都是非凸的，因此非凸优化(Non-Convex Optimization)受到越来越多的重视。

如何判断问题是不是非凸问题呢？若可行域为非凸集合，则为非凸优化问题，例如（混合）整数规划问题；若目标函数是非凸函数，亦为非凸问题。Stephen Boyd (https://en.wikipedia.org/wiki/Stephen_P._Boyd) 在凸规划中指出$x^a, 0 \le a \le 1$和$\log x$为凹函数。下面是一些凹目标函数的具体问题：

- 可靠最短路问题(Xing and Zhou, 2011)：为了表达路段时间的随机性，既考虑出行时间的均值又考虑方差，目标函数为最小化路径出行时间均值和标准差的线性组合。

  $\min\{mean + \beta\sqrt{var}\}$，$mean$为路径出行时间均值，属于线性部分；$\sqrt{var}$为路径出行时间标准差，属于非线性、凹函数部分。$\beta$为可靠度系数，与出行者的风险态度相关，下面对其进行进一步解释。假设路段出行时间服从正态分布且无路段关联，根据出行者期望的准时到达概率α，可将出行者分为三类：风险中立型出行者(Risk-Neutral，$\alpha = 0.5$)，则$\beta = 0$，此时最可靠路径问题等价于最小期望路劲问题；风险规避型出行者(Risk-Averse，$\alpha > 0.5$)，则$\beta > 0$；风险追逐型出行者(Risk-Seeking，$\alpha < 0.5$)，则$\beta < 0$。一般认为出行者是风险规避型，β一种典型值为 1.27。
- 选址-库存联合问题(Shen and Coullard, Daski, 2003)：有供货商和多个零售商，零售商对商品的需求具有不确定性。为了对抗不确定性，部分零售商同时作为配送中心，以提高服务水平。问题是如何确定这些配送中心及配送中心服务哪些零售商。

  目标函数是最小化选址成本+运输成本+库存成本+可靠性成本：

  $\min\{fix\_cost + trans\_cost + inven\_cost + \theta\sqrt{invent\_var}\}$。$inven_var$为所有零售商的商品需求量方差之和，$\sqrt{invent_var}$为非线性、凹函数项；其余各项均为线性部分。
- 路段容量约束下考虑经济规模效应(Economy of Scales)的最小凹成本网络流问题(Larsson et al., 1994)：路段成本函数为凹函数；考虑路段的拥堵效应；考虑经济规模效应。这里对拥挤(Crowding)和拥堵(Congestion)进行简单解释：车辆在拥挤路段花费时间全部为车辆行驶时间；车辆在拥堵路段花费时间包括车辆行驶时间和停车等待时间。

  $\min \sum_{ij \in A} f_{ij}(x_{ij})$，路段$ij$的成本函数$f_{ij}\left(x_{ij}\right)$是凹函数，则目标函数为凹函数。

那么如何求解非凸优化问题呢？由于非凸优化问题的可行域内存在无数局部最优点，通常得到全局最优解的算法复杂度是指数级别的。最经典的算法是蒙特卡洛投点法（Monte Carlo Algorithm）：在可行域内随机投一个点进行局部搜索得到局部最优解，直到满足终止条件。下文对可靠最短路问题给出拉格朗日替代法（Lagrangian Substitution Method）的解决方案：通过变量分离方法（Variable-

Splitting Method）和拉格朗日分解方法（Lagrangian Decomposition Method），将松弛问题分解为易求解的子问题；松弛问题的目标值作为原问题最优值的下界，原问题可行解代入原问题作为原问题最优值的上界；利用次梯度法更新乘子，在迭代过程不断更新上下界，进而得到原问题的（近似）全局最优解。

最短路问题是运筹学中一个基本问题，经常作为求解复杂问题的子问题。在交通网络中指起讫点之间路段成本和（下文以出行时间为例）最小的路径。在标准的最短路问题中，路段出行时间已知并且固定不变。但在实际的交通网络中由于不确定因素（如恶劣天气，事故等）会导致交通供给和需求的不确定性，最终使得路段出行时间呈现出随机性。图 1 是使用实际观测数据绘制的某路段出行时间频率分布图。忽略出行时间的随机性可能给出行者带来较大的影响，例如去参加面试或会议，迟到会给出行者带来很大的损失。因此考虑出行时间随机性是十分必要的，并且将这类最短路问题称为可靠最短路问题。图 2 中为最短路问题和可靠最短路问题的简单例子。

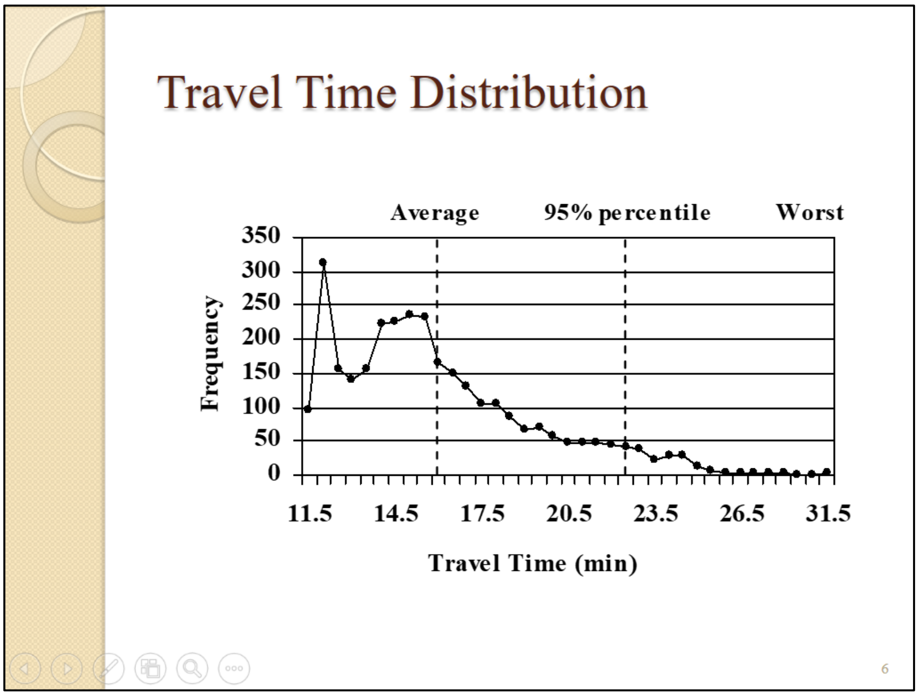

图 1 某路段出行时间分布频率

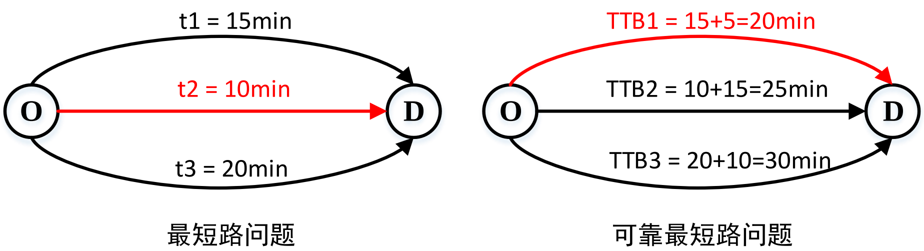

图 2 最短路问题和可靠最短路问题

可靠最短路问题根据可靠度可以分为出行时间可靠度和到达时间可靠度。出行时间可靠度指出行过程的可靠性，分为最大化准时到达概率（已知出行时间阈值）和最小化百分位出行时间（已知出行时间分位数）。出行时间可靠度定义考虑出行时间的鲁棒性(Robustness)，详见图 4。实际生活中，出行者往往会选择出行时间均值较小的路径，并且预留部分时间来应对出行过程中的随机性，到达时间可靠度可以比较准确的描绘这一现象。到达时间可靠度指出行者准时到达目的地的可靠性，目标函数一般是最小化出行时间预算（TTB=期望出行时间+预留出行时间）。到达时间可靠度定义考虑出行时间的不确定性(Variability)，详见图 3。另外，路网中的路段出行时间是相互关联的：考虑路段关联即考虑协方差信息会使得求出的结果更贴近现实，缺点是计算比较复杂；不考虑路段关联必然会丢失一些信息，但模型简单容易求解。

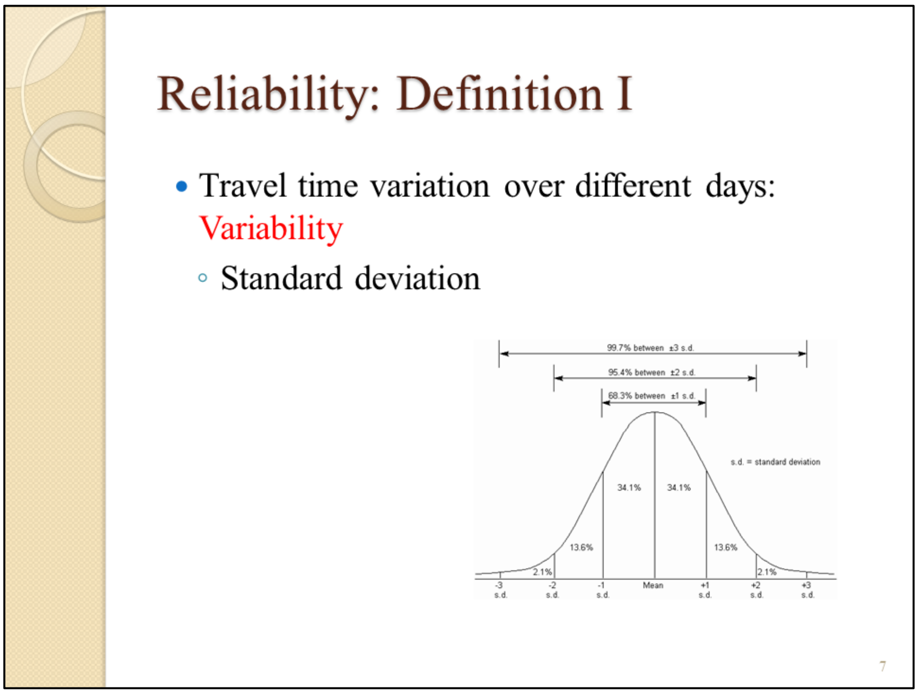

图 3 出行时间可靠度定义I

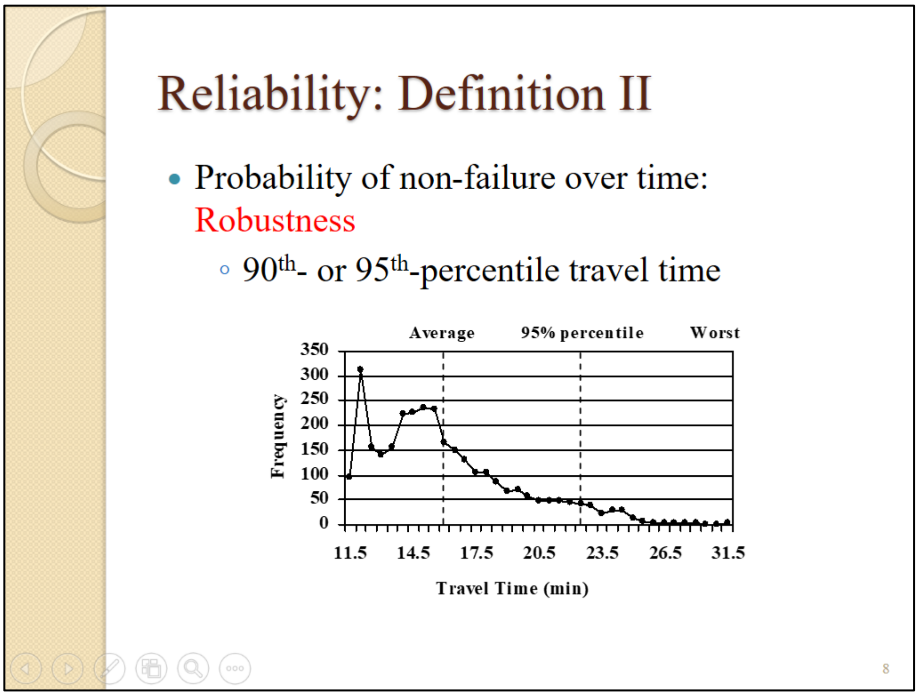

图 4 出行时间可靠度定义II

我们下面介绍可靠度模型的目标函数采用了到达时间可靠度的形式，且不考虑路段出行时间关联，该问题的一种解决方案(Xing and Zhou, 2011)如下：

1. 该文引入 0-1 变量表示路段选择，将可靠最短路问题建模为整数非线性问题，即图 5 中的 Model I。

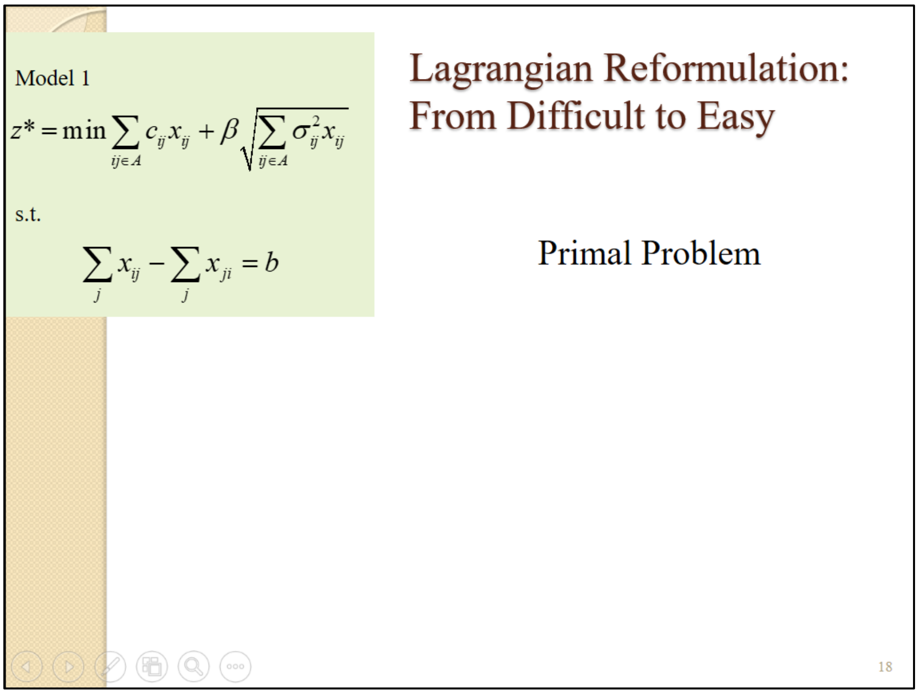

图 5 整数非线性模型

其中A代表路段集合，ij代表路段，c_{ij}为路段时间均值，σ_{ij}^2为路段时间方差，β为可靠性系数与出行者的风险态度相关，x_{ij}为 0-1 决策变量。

接下来对此模型进行进一步解释：

经济学中一类投资优化问题——多目标投资优化(Multi-Objective Portfolio Optimization)：在满足约束、利用历史信息的条件下，优化多个变量构成的目标量(Quantity)。目标量在经济学中叫做夏普指数(Sharpe Ratio)，用来权衡投资中的收益和风险。夏普指数的公式是$Sharpe\ \ Ratio = mean/standard\ \ deviation$。最大化夏普指数意味着同时优化收益和风险——最大化收益、最小化风险。显然最大化夏普系数对应的投资者为风险厌恶类型。

而我们的优化问题需要同时优化均值和标准差(Simultaneous Optimization of Mean and Standard Deviation)，故问题也属于多目标优化问题。显然优化夏普系数不适用于我们的问题：想要缩减出行时间均值，又想尽可能的减小风险。因此这里采用了均值和标准差的线性组合，并且系数的调整可以对应多种风险态度的出行者，那么$\mathrm{Ratio} = mean + \beta\sqrt{var}$。

想要进一步了解 Multi-Objective Portfolio Optimization 的读者可以参考：

- Markowitz’s Mean-Variance Analysis
- https://en.wikipedia.org/wiki/Modern_portfolio_theory
- Mean Absolute Deviation Optimization (MAD)
- https://www.tandfonline.com/doi/abs/10.1080/08982110903394205?journalCode=lqen20

2. 使用拉格朗日松弛方法将难约束吸纳到目标函数中是求解整数规划问题的重要方法。在此基础上，学者们提出了变量分离方法和拉格朗日分解方法：引入辅

助变量，将原变量分离成一对变量，变量对需要满足一致性约束；利用拉格朗日松弛方法将一致性约束吸纳到目标函数中并进行问题分解。例如 Fisher 老师(1997)使用变量分离方法将 VRP 问题中的服务满足约束分解为一致性约束和服务满足约束（形式不同），将一致性约束吸纳到目标函数中，并将其分解为易求解的最短路子问题和半分配子问题。

3. 为了处理目标函数中的非线性和非凸性，引入辅助变量替换路径出行时间方差部分，即$y=\sum_{ij\in A}\sigma_{ij}^2 x_{ij}$。我们称$y=\sum_{ij\in A}\sigma_{ij}^2 x_{ij}$为一致性约束，可将其松弛为$y\geq\sum_{ij\in A}\sigma_{ij}^2 x_{ij}$，松弛后的约束是紧的（tight or active），即约束松弛后不会不改变问题的最优解。可靠最短路问题得到的路径时间均值必然大于等于最小期望路径得到的路径时间均值，即可靠最短路问题的解在图 6 上侧部分。假设可靠最短路问题的方差大于最小期望路径的方差y'，即$y\geq y'$，意味着可靠最短路的可靠度目标值大于等于最小期望路径对应的可靠度目标值。即最小期望路径的可靠度大于等于最可靠路径的可靠度，显然这与事实不符，所以可靠最短路问题的解在图 6 左侧部分，即$0\leq y\leq y'$。则模型变化为图 7 中的 Model II：

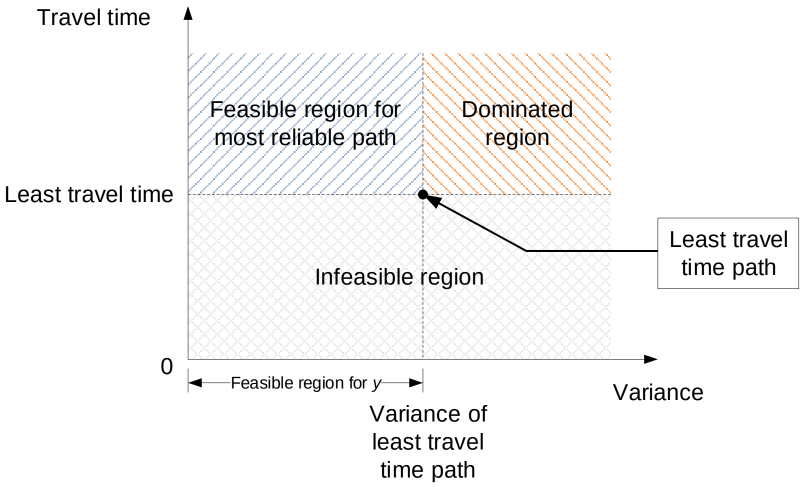

图 6 新增变量y可行范围示意图

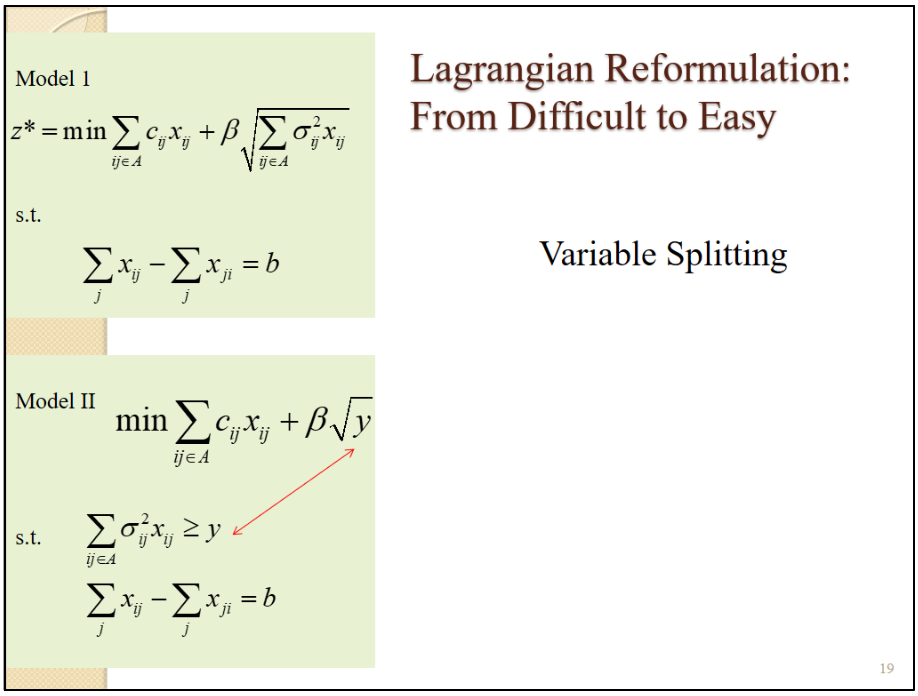

图 7 变量分离方法

4. 由于处理后的模型依旧难以求解，采用拉格朗日松弛方法将松弛后的一致性约束$y \geq \sum_{ij\in A}\sigma_{ij}^2 x_{ij}$吸纳到目标函数中，得到拉格朗日松弛问题，见图 8 中的 Model III。松弛问题的可行域较原问题变大，目标函数值较原问题变小，但问题的复杂度降低。显然拉格朗日松弛问题的目标值是原问题最优值的下界，但我们需要的是找到最接近原问题最优值的下界，即拉格朗日对偶问题$L^* = \max_{\mu\geq 0} L(\mu)$。

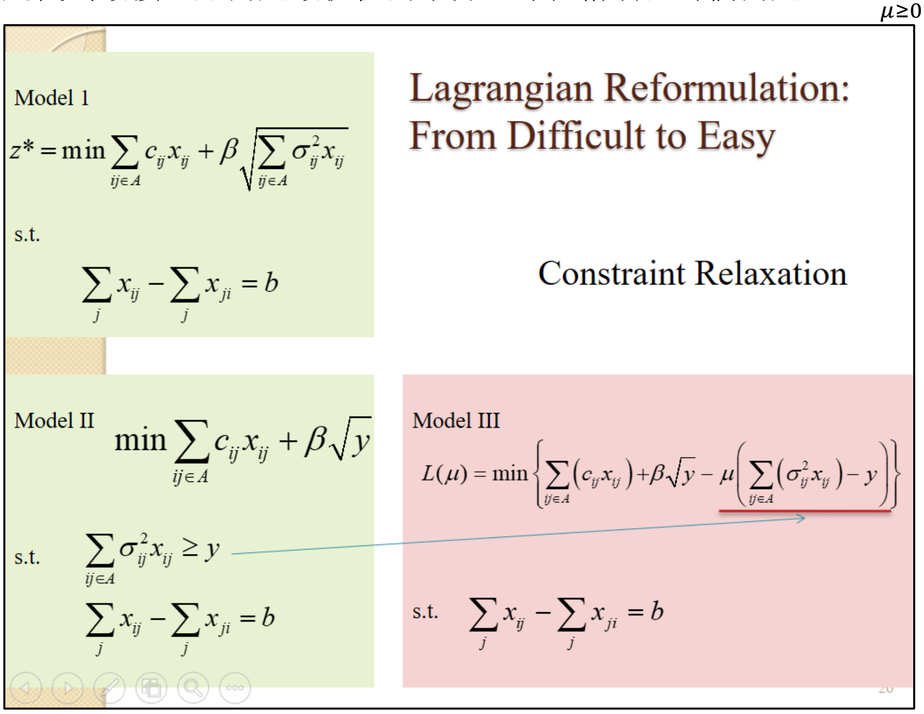

图 8 拉格朗日松弛

5. 接下来对拉格朗日松弛问题按照原变量和辅助变量进行重新分组，可对松弛问题进行分解：子问题 1，广义成本最短路问题，可以采用传统的最短路算法求

解；子问题 2，关于辅助变量y的单变量凹函数最小值问题，极值点在端点取得。拉格朗日分解过程详见图 8。子问题求解完成即可获得拉格朗日松弛问题目标值，即本次迭代获得的下界$L_k(\mu_k)$。进一步更新全局下界，$max\{L(\mu),LB_{k-1}\} \rightarrow LB_k$。

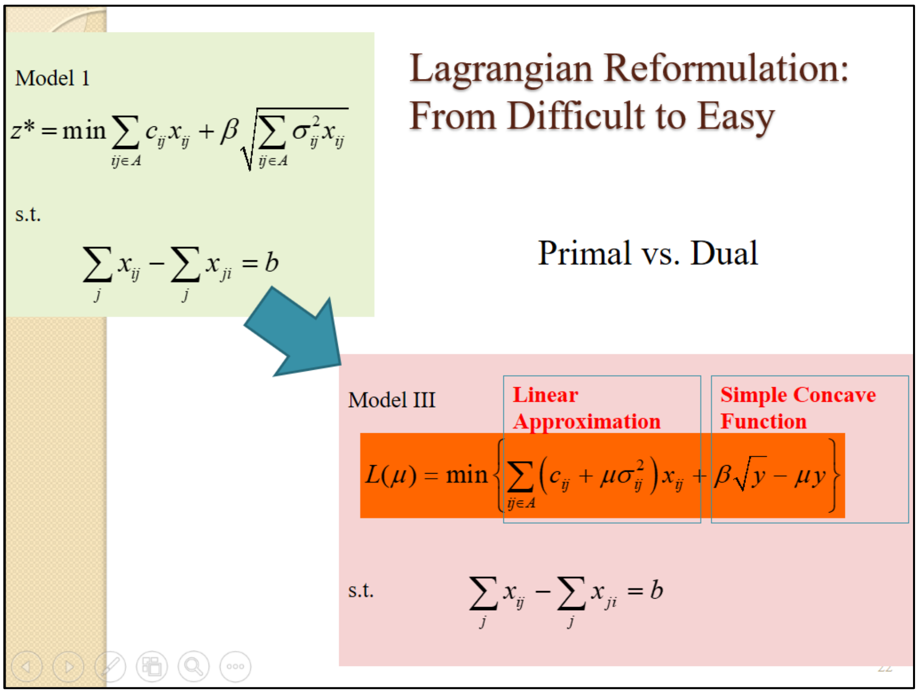

图 8 拉格朗日分解

6. 仅仅得到原问题的下界是不够的，此时我们无法评估得到的下界与原问题最优值的偏差。因此，我们还需要得到原问题最优值的上界。显然，原问题的任意可行解代入到原问题都可以作为最优值的上界。又考虑到子问题 1 为广义成本最短路问题，对子问题 1 的求解结果进行处理（使其满足原问题的约束）后即可得到原问题可行解，上下界更新示意图见图 9。采用次梯度法更新拉格朗日乘子，次梯度法求解过程见图 10。迭代更新上下界的值，上下界差值小于预先设置的可接受偏差即可停止迭代，即可得到原问题的（近似）全局最优解。

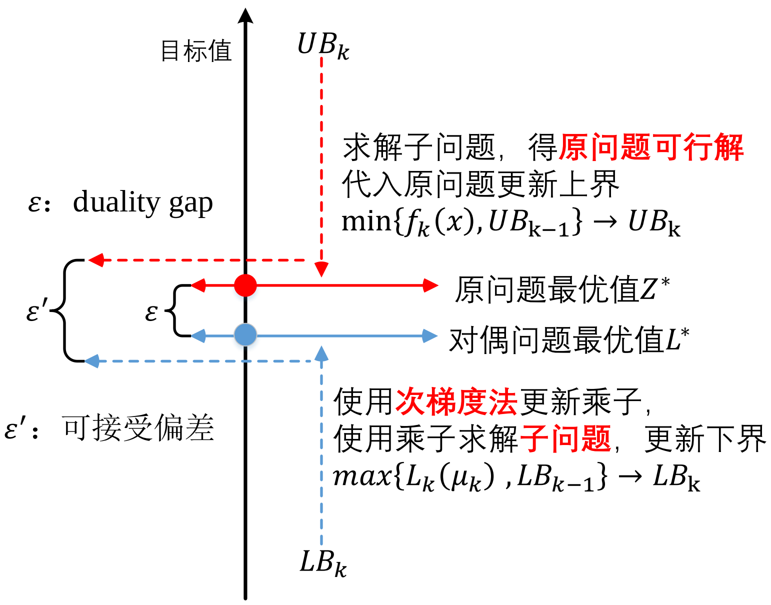

图 9 上下界更新示意图

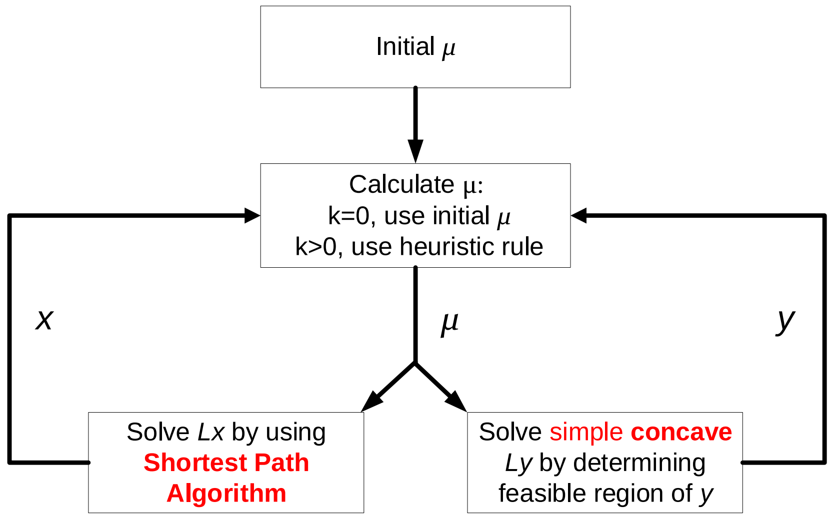

图 10 次梯度法求解过程

7. 荷兰学者 Doek 对以上算法进行了测试，初始化参数：可接受偏差$\varepsilon' = 10^{-1}$，迭代次数$K = 50$，可靠性系数$\beta = 1.27$。测试结果：在 1,000 个结点的网络中，平均相对误差为1.0%~2.0%，运行时间为0.096s~0.108s；在 10,000 个结点的网络中，平均相对误差为0.9%~4.8%，运行时间为2.914s~3.606s。由于方差大小会影响测试结果，测试中采用方差较小和方差较大两个场景，所以测试结果采用区间形式给出。测试结果表示这种可靠路径问题的解决方案在大型网络中可以在可接受的时间内得到比较满意的最优解。另外，使用 python 平台编写得简单算例可在 https://github.com/songmaocan/Reliable_Shortest_Path_problem 下载。

参考文献：

Doek, G., 2018. Finding the most reliable path in a stochastic network.

Fisher, M.L., Jörnsten, K.O., Madsen, O.B., 1997. Vehicle routing with time windows: Two optimization algorithms. *Operations research* 45, 488-492.

Larsson, T.r., Migdalas, A., R6nnqvist, M., 1994. A Lagrangean heuristic for the capacitated concave minimum cost network flow problem. *European Journal of Operational Research* 78, 116-129.

Shen, Z.-J.M., Coullard, C., Daskin, M.S., 2003. A Joint Location-Inventory Model. *Transportation Science* 37, 40-55.

Xing, T., Zhou, X.S., 2011. Finding the most reliable path with and without link travel time correlation: A Lagrangian substitution based approach. *Transport Res B-Meth* 45, 1660-1679.