Gradient Methods on Network Flow Problems

The network flow problem we consider is

$$\begin{array}{ll} \underset{x=\{x_{ij}\}_{(i,j)\in\mathcal{E}}}{\text{minimize}} & \sum_{(i,j)\in\mathcal{E}} \phi_{ij}(x_{ij}) \\ \text{subject to} & Bx = d \\ & 0 \le x \le c\,, \end{array} \tag{1}$$

which is associated to a digraph with P nodes and E edges (the set of edges is $\mathcal{E}$). That digraph is represented with its node-arc incidence matrix $B \in \mathbb{R}^{P\times E}$ and each edge (i,j) in the network has capacity c_{ij}. d is the vector of flow inputs/outputs, and ϕ_{ij} is the cost associated to edge (i,j). We assume that $(i,j)\in\mathcal{E}$ means a connection (arrow) from i to j. Also, the ijth column of the node-arc incidence matrix has a -1 in the ith entry and a 1 in the jth entry. Consider the dual of (1):

$$\underset{\lambda=(\lambda_1,\ldots,\lambda_P)}{\text{maximize}} \quad L(\lambda)\,, \tag{2}$$

where the dual function is

$$\begin{aligned} L(\lambda) &= \inf_{0\le x\le c} \sum_{(i,j)\in\mathcal{E}} \phi_{ij}(x_{ij}) + \left(B^\top\lambda\right)^\top x - d^\top\lambda \\ &= -d^\top\lambda + \sum_{(i,j)\in\mathcal{E}} \inf_{0\le x_{ij}\le c_{ij}} \phi_{ij}(x_{ij}) + (\lambda_j-\lambda_i)x_{ij} \\ &= -d^\top\lambda + \sum_{(i,j)\in\mathcal{E}} \phi_{ij}(x_{ij}(\lambda_j-\lambda_i)) + (\lambda_j-\lambda_i)x_{ij}(\lambda_j-\lambda_i)\,, \end{aligned}$$

where

$$x_{ij}(v) := \inf_{0\le x_{ij}\le c_{ij}} \phi_{ij}(x_{ij}) + v\,x_{ij}\,, \tag{3}$$

which is always well-defined. We will choose

$$\phi_{ij}(x_{ij}) = \frac{x_{ij}}{c_{ij}-x_{ij}}\,,$$

which is continuously differentiable and strictly convex in $0 \le x_{ij} < c_{ij}$. To find (3) for this choice, take the derivative of the objective of (3) and equate it to zero:

$$\begin{aligned} & \frac{c_{ij}}{(c_{ij}-x_{ij})^2} + v = 0 \\ \iff\quad & c_{ij} + v(c_{ij}-x_{ij})^2 = 0 \\ \iff\quad & v x_{ij}^2 - 2vc_{ij}x_{ij} + c_{ij} + vc_{ij}^2 = 0 \\ \iff\quad & x_{ij}^2 - 2c_{ij}x_{ij} + \frac{c_{ij}+vc_{ij}^2}{v} = 0\,, \end{aligned}$$

whose solutions, if they exist, are

$$\begin{aligned} x_{ij} &= \frac{2c_{ij} \pm \sqrt{4c_{ij}^2 - 4c_{ij}^2 - 4c_{ij}/v}}{2} \\ &= \frac{2c_{ij} \pm \sqrt{-4c_{ij}/v}}{2} \\ &= c_{ij} - \sqrt{-c_{ij}/v}\,, \end{aligned}$$

where we selected the solution with $-$ because we need $x_{ij} \leq c_{ij}$. This quadratic form has a solution only if $v < 0$. In fact, it can be seen graphically that if $v \geq 0$, $x_{ij}(v) = 0$. Therefore,

$$x_{ij}(v) = \begin{cases} 0 & v \geq 0 \\ \left[c_{ij} - \sqrt{-c_{ij}/v}\right]_{[0,c_{ij}]} & v < 0\,, \end{cases}$$

where $[\cdot]_{[0,c_{ij}]}$ denotes the projection onto the set $[0, c_{ij}]$.

The gradient of $L(\lambda)$ is given by $\nabla L(\lambda) = Bx(\lambda) - d$, whose pth component is

$$\frac{\partial}{\partial \lambda_p} L(\lambda) = -\sum_{j \in \mathcal{N}_p} x_{pj}(\lambda_j - \lambda_p) + \sum_{j \in \mathcal{N}_p} x_{jp}(\lambda_p - \lambda_j) - d_p\,.$$

Therefore, λ_p can be updated at node p through

$$\lambda_p^{k+1} = \lambda_p^k + \alpha\Big(\sum_{j \in \mathcal{N}_p} x_{jp}(\lambda_p^k - \lambda_j^k) - \sum_{j \in \mathcal{N}_p} x_{pj}(\lambda_j^k - \lambda_p^k) - d_p\Big)\,,$$

where α is the stepsize and the algorithm is the simple gradient method. The previous update is possible if both nodes i and j know ϕ_{ij} and all the nodes exchange their λ^k's after updating them. Note that every computation involving ϕ_{ij} takes place at both nodes i and j. Note also that $\nabla L(\lambda)$ is not Lipschitz continuous.