

Fast Gradient Methods for Network Flow Problems

Yurii Nesterov, CORE/INMA (UCL)

May 27-29, 2009, **IP at CORE**, UCL
(In honor of L.A. Wolsey)

Anstreicher, K.M.,Wolsey, L.A.: On dual solutions in subgradient optimization. Center for Operations Research and Econometrics. Louvain-la-Neuve, Belgium (**1992**, working paper)

Anstreicher, K.M., Wolsey, L.A.: On dual solutions in subgradient optimization. Center for Operations Research and Econometrics. Louvain-la-Neuve, Belgium (**1992**, working paper)

Published in: Mathematical Programming, August 2009, 213-220.

Outline

- 1 Bounds for optimal network arc flows
- 2 Maximal concurrent flow problem
- 3 Minimization in relative scale
- 4 Subgradient method
- 5 Smoothing technique
- 6 Conclusion

Optimization problems over the set of bounded arc flows

Notation

- $\mathcal{N} = \{1, \dots, n\}$ - set of nodes.
- $\mathcal{A} = \{1, \dots, m\}$ - set of directed arcs.
- $\delta_\alpha = e_{h(\alpha)} - e_{t(\alpha)} \in R^n$, ($h(\alpha)$, $t(\alpha)$): the head and tail of α)
- $E = (\delta_1, \dots, \delta_m) \in R^{n \times m}$ is the *balance matrix*. $E^T e = 0$.
- $\mathcal{F}(d) = \{f \in R^m : 0 \leq f \leq \bar{f}, Ef = d\}$ is the set of *feasible flows* satisfying the demands $d \in R^n$.

Optimization problems over the set of bounded arc flows

Notation

- $\mathcal{N} = \{1, \dots, n\}$ - set of nodes.
- $\mathcal{A} = \{1, \dots, m\}$ - set of directed arcs.
- $\delta_\alpha = e_{h(\alpha)} - e_{t(\alpha)} \in R^n$, ($h(\alpha)$, $t(\alpha)$): the head and tail of α)
- $E = (\delta_1, \dots, \delta_m) \in R^{n \times m}$ is the *balance matrix*. $E^T e = 0$.
- $\mathcal{F}(d) = \{f \in R^m : 0 \leq f \leq \bar{f}, Ef = d\}$ is the set of *feasible flows* satisfying the demands $d \in R^n$.

Main question: how to control the size of set \mathcal{F}

Optimization problems over the set of bounded arc flows

Notation

- $\mathcal{N} = \{1, \dots, n\}$ - set of nodes.
- $\mathcal{A} = \{1, \dots, m\}$ - set of directed arcs.
- $\delta_\alpha = e_{h(\alpha)} - e_{t(\alpha)} \in R^n$, ($h(\alpha)$, $t(\alpha)$): the head and tail of α)
- $E = (\delta_1, \dots, \delta_m) \in R^{n \times m}$ is the *balance matrix*. $E^T e = 0$.
- $\mathcal{F}(d) = \{f \in R^m : 0 \leq f \leq \bar{f}, Ef = d\}$ is the set of *feasible flows* satisfying the demands $d \in R^n$.

Main question: how to control the size of set \mathcal{F}

- By demands? (Circulations: $f \geq 0$, $Ef = 0$.)

Optimization problems over the set of bounded arc flows

Notation

- $\mathcal{N} = \{1, \dots, n\}$ - set of nodes.
- $\mathcal{A} = \{1, \dots, m\}$ - set of directed arcs.
- $\delta_\alpha = e_{h(\alpha)} - e_{t(\alpha)} \in R^n$, ($h(\alpha)$, $t(\alpha)$): the head and tail of α)
- $E = (\delta_1, \dots, \delta_m) \in R^{n \times m}$ is the *balance matrix*. $E^T e = 0$.
- $\mathcal{F}(d) = \{f \in R^m : 0 \leq f \leq \bar{f}, Ef = d\}$ is the set of *feasible flows* satisfying the demands $d \in R^n$.

Main question: how to control the size of set \mathcal{F}

- By demands? (Circulations: $f \geq 0$, $Ef = 0$.)
- By arc flow bounds $\bar{f}^{(\alpha)}$? (Can be big or infinite).

Acyclic optimal solutions

In some situations, the optimal flow $f^* \in \mathcal{F}(d)$ contains no cycles:

$$f^* \neq f_1 + f_2, \quad f_1 \in \mathcal{F}(d), \quad f_2 \geq 0, \quad \underbrace{Ef_2 = 0}_{\text{circulation}}, \quad f_2 \neq 0.$$

Acyclic optimal solutions

In some situations, the optimal flow $f^* \in \mathcal{F}(d)$ contains no cycles:

$$f^* \neq f_1 + f_2, \quad f_1 \in \mathcal{F}(d), \quad f_2 \geq 0, \quad \underbrace{Ef_2 = 0}_{\text{circulation}}, \quad f_2 \neq 0.$$

Examples

- Objective function depends only on demands.

Acyclic optimal solutions

In some situations, the optimal flow $f^* \in \mathcal{F}(d)$ contains no cycles:

$$f^* \neq f_1 + f_2, \quad f_1 \in \mathcal{F}(d), \quad f_2 \geq 0, \quad \underbrace{Ef_2 = 0}_{\text{circulation}}, \quad f_2 \neq 0.$$

Examples

- Objective function depends only on demands.
- Objective function is properly monotone in arc flows.

Acyclic optimal solutions

In some situations, the optimal flow $f^* \in \mathcal{F}(d)$ contains no cycles:

$$f^* \neq f_1 + f_2, \quad f_1 \in \mathcal{F}(d), \quad f_2 \geq 0, \quad \underbrace{Ef_2 = 0}_{\text{circulation}}, \quad f_2 \neq 0.$$

Examples

- Objective function depends only on demands.
- Objective function is properly monotone in arc flows.

Important:

The volume of acyclic flow can be bounded by maximal node-to-node flows.

Maximal flows

Definition: for $i, j \in \mathcal{N}$, $i \neq j$, denote

$$\theta_{i,j}^* = \max_{\tau, f} \{ \tau : Ef = \tau(e_j - e_i), 0 \leq f \leq \bar{f} \},$$

Maximal flows

Definition: for $i, j \in \mathcal{N}$, $i \neq j$, denote

$$\theta_{i,j}^* = \max_{\tau, f} \{ \tau : Ef = \tau(e_j - e_i), 0 \leq f \leq \bar{f} \}, \quad \theta_{i,i}^* = \sum_{j \neq i} \theta_{i,j}^*.$$

Maximal flows

Definition: for $i, j \in \mathcal{N}$, $i \neq j$, denote

$$\theta_{i,j}^* = \max_{\tau, f} \{ \tau : Ef = \tau(e_j - e_i), 0 \leq f \leq \bar{f} \}, \quad \theta_{i,i}^* = \sum_{j \neq i} \theta_{i,j}^*.$$

Denote $\mathcal{D}_k = \{d \in R^n : d^{(i)} \geq 0, i \neq k, \langle e, d \rangle = 0\}$.

Lemma: Let $f \in \mathcal{F}(d_k)$ for some $d_k \in \mathcal{D}_k$. If f is acyclic, then

$$\max \left\{ \sum_{\alpha: h(\alpha)=i} f^{(\alpha)}, \sum_{\alpha: t(\alpha)=i} f^{(\alpha)} \right\} \leq \theta_{k,i}^*.$$

Bounds for acyclic multicommodity flows

Denote by $\mathcal{S} = \{1, \dots, q\}$ the set of *commodity sources*, and

$$\Theta_*^{(i,k)} = \min\{\theta_{k,h(i)}^*, \theta_{k,t(i)}^*\}, \quad i \in \mathcal{A}, k \in \mathcal{S}.$$

Thus, $\Theta_* \in R^{m,q}$ is a matrix.

Bounds for acyclic multicommodity flows

Denote by $\mathcal{S} = \{1, \dots, q\}$ the set of *commodity sources*, and

$$\Theta_*^{(i,k)} = \min\{\theta_{k,h(i)}^*, \theta_{k,t(i)}^*\}, \quad i \in \mathcal{A}, k \in \mathcal{S}.$$

Thus, $\Theta_* \in R^{m,q}$ is a matrix.

Assume that all flows $f_k \in \mathcal{F}(d_k)$ with $d_k \in \mathcal{D}_k$ are acyclic.

Then, for flow matrix $F = (f_1, \dots, f_q)$ we have

$$0 \leq F \leq \Theta_*, \quad F \cdot e \leq \bar{f}.$$

Bounds for acyclic multicommodity flows

Denote by $\mathcal{S} = \{1, \dots, q\}$ the set of *commodity sources*, and

$$\Theta_*^{(i,k)} = \min\{\theta_{k,h(i)}^*, \theta_{k,t(i)}^*\}, \quad i \in \mathcal{A}, k \in \mathcal{S}.$$

Thus, $\Theta_* \in R^{m,q}$ is a matrix.

Assume that all flows $f_k \in \mathcal{F}(d_k)$ with $d_k \in \mathcal{D}_k$ are acyclic.

Then, for flow matrix $F = (f_1, \dots, f_q)$ we have

$$0 \leq F \leq \Theta_*, \quad F \cdot e \leq \bar{f}.$$

Note: 1. Matrix Θ_* can be computed in $O(n^2 m \ln n)$ operations.

Bounds for acyclic multicommodity flows

Denote by $\mathcal{S} = \{1, \dots, q\}$ the set of *commodity sources*, and

$$\Theta_*^{(i,k)} = \min\{\theta_{k,h(i)}^*, \theta_{k,t(i)}^*\}, \quad i \in \mathcal{A}, k \in \mathcal{S}.$$

Thus, $\Theta_* \in R^{m,q}$ is a matrix.

Assume that all flows $f_k \in \mathcal{F}(d_k)$ with $d_k \in \mathcal{D}_k$ are acyclic.

Then, for flow matrix $F = (f_1, \dots, f_q)$ we have

$$0 \leq F \leq \Theta_*, \quad F \cdot e \leq \bar{f}.$$

- Note:** 1. Matrix Θ_* can be computed in $O(n^2 m \ln n)$ operations.
2. The bounds are row-separable.

Maximal concurrent flow problem (MCFP)

Let us fix demands $d_k \in \mathcal{D}_k$, $k = 1, \dots, q$. Our goal is to find

$$\tau^* = \max_{\tau, F = \{f_k\}_{k \in \mathcal{S}}} \{ \tau : Ef_k \geq \tau \cdot d_k, k \in \mathcal{S}, \\ Fe \leq \bar{f}, F \geq 0 \}.$$

Maximal concurrent flow problem (MCFP)

Let us fix demands $d_k \in \mathcal{D}_k$, $k = 1, \dots, q$. Our goal is to find

$$\tau^* = \max_{\tau, F = \{f_k\}_{k \in \mathcal{S}}} \left\{ \tau : E f_k \geq \tau \cdot d_k, k \in \mathcal{S}, \right. \\ \left. F e \leq \bar{f}, F \geq 0 \right\}.$$

Dualizing the demand constraints, we get

$$\begin{aligned} \tau^* &= \max_{\substack{F \geq 0, \tau \in \mathbb{R}, \\ F e \leq \bar{f}}} \min_{\{T_k \geq 0\}} \left\{ \tau + \sum_{k=1}^q \langle T_k, E f_k - \tau \cdot d_k \rangle \right\} \\ &= \min_{\{T_k \geq 0\}} \left[\max_{\substack{F \geq 0, \\ F e \leq \bar{f}}} \left\{ \sum_{k=1}^q \langle T_k, E f_k \rangle \right\} : \sum_{k=1}^q \langle d_k, T_k \rangle = 1 \right] \\ &= \min_{\{T_k \geq 0\}} \left[\sum_{j=1}^m \bar{f}^{(j)} \left(\max_{1 \leq k \leq q} \langle \delta_j, T_k \rangle \right) + \sum_{k=1}^q \langle d_k, T_k \rangle = 1 \right]. \end{aligned}$$

Dual problem

Denote $C_k = \{T_k \in R_+^n : T_k^{(k)} = 0\}$, $T = (T_1, \dots, T_q)$, and

$$\phi(T) = \sum_{j=1}^m \bar{f}^{(j)} \max_{1 \leq k \leq q} \langle \delta_j, T_k \rangle_+.$$

Then the dual MCFP is as follows

$$\text{Find } \tau^* = \min_T \{ \phi(T) : \sum_{k=1}^q \langle d_k, T_k \rangle = 1, T_k \in C_k, 1 \leq k \leq q \}. \quad (*)$$

Dual problem

Denote $C_k = \{T_k \in R_+^n : T_k^{(k)} = 0\}$, $T = (T_1, \dots, T_q)$, and

$$\phi(T) = \sum_{j=1}^m \bar{f}^{(j)} \max_{1 \leq k \leq q} \langle \delta_j, T_k \rangle_+.$$

Then the dual MCFP is as follows

$$\text{Find } \tau^* = \min_T \{ \phi(T) : \sum_{k=1}^q \langle d_k, T_k \rangle = 1, T_k \in C_k, 1 \leq k \leq q \}. \quad (*)$$

Note:

- Complexity of (*) depends on the Lipschitz constant for ϕ .
- This Lipschitz constant depends on \bar{f} .
- For some arcs, these bounds can be arbitrary big.

Dual problem

Denote $C_k = \{T_k \in R_+^n : T_k^{(k)} = 0\}$, $T = (T_1, \dots, T_q)$, and

$$\phi(T) = \sum_{j=1}^m \bar{f}^{(j)} \max_{1 \leq k \leq q} \langle \delta_j, T_k \rangle_+.$$

Then the dual MCFP is as follows

$$\text{Find } \tau^* = \min_T \{ \phi(T) : \sum_{k=1}^q \langle d_k, T_k \rangle = 1, T_k \in C_k, 1 \leq k \leq q \}. \quad (*)$$

Note:

- Complexity of (*) depends on the Lipschitz constant for ϕ .
- This Lipschitz constant depends on \bar{f} .
- For some arcs, these bounds can be arbitrary big.

We need reasonable bounds on arc flows.

Regularized dual problem

Main Idea: The optimal flows in MCFP are acyclic \Rightarrow Apply the max-flow bounds.

Regularized dual problem

Main Idea: The optimal flows in MCFP are acyclic \Rightarrow Apply the max-flow bounds.

Regularized dual function

$$\tilde{\phi}(T) = \max_F \left\{ \sum_{k=1}^q \langle T_k, E f_k \rangle : 0 \leq F \leq \Theta_*, F e \leq \bar{f} \right\}$$

with $T \in \mathcal{R}_S \stackrel{\text{def}}{=} \{T \in R^{n \times q} : T_k^{(k)} = 0, k = 1, \dots, q\}$.
Can be computed in $O(mq \ln q)$ operations.

Regularized dual problem

Main Idea: The optimal flows in MCFP are acyclic \Rightarrow Apply the max-flow bounds.

Regularized dual function

$$\tilde{\phi}(T) = \max_F \left\{ \sum_{k=1}^q \langle T_k, Ef_k \rangle : 0 \leq F \leq \Theta_*, Fe \leq \bar{f} \right\}$$

with $T \in \mathcal{R}_S \stackrel{\text{def}}{=} \{T \in R^{n \times q} : T_k^{(k)} = 0, k = 1, \dots, q\}$.

Can be computed in $O(mq \ln q)$ operations.

New dual problem: $\tau^* = \min_{T \in \mathcal{R}_S} \{\tilde{\phi}(T) : \sum_{k=1}^q \langle d_k, T_k \rangle = 1\}$

has the same optimal solution as (*).

Note: The subspace \mathcal{R}_S can be replaced by the cone $\mathcal{C} \stackrel{\text{def}}{=} \prod_{k=1}^q \mathcal{C}_k$.

Minimization in relative scale

Problem formulation:

$$\text{Find } f^* = \min_x f(x) : Ax = b.$$

where $b \neq 0$ and

- f is homogeneous: $f(x) = \max_{x \in Q} \langle s, x \rangle$, $0 \in \text{int } Q$.

Assumption: $B(\gamma_1) = \{x : \|x\| \leq \gamma_1\} \subseteq Q \subseteq B(\gamma_2)$. Then

$$\gamma_1 \|x\| \leq f(x) \leq \gamma_2 \|x\|.$$

Minimization in relative scale

Problem formulation:

$$\text{Find } f^* = \min_x f(x) : Ax = b.$$

where $b \neq 0$ and

- f is homogeneous: $f(x) = \max_{x \in Q} \langle s, x \rangle$, $0 \in \text{int } Q$.

Assumption: $B(\gamma_1) = \{x : \|x\| \leq \gamma_1\} \subseteq Q \subseteq B(\gamma_2)$. Then

$$\gamma_1 \|x\| \leq f(x) \leq \gamma_2 \|x\|.$$

Subgradient method: $f(x_k) - f^* \leq \frac{L \|x_0 - x^*\|}{\sqrt{k+1}}$.

Minimization in relative scale

Problem formulation:

$$\text{Find } f^* = \min_x f(x) : Ax = b.$$

where $b \neq 0$ and

- f is homogeneous: $f(x) = \max_{x \in Q} \langle s, x \rangle$, $0 \in \text{int } Q$.

Assumption: $B(\gamma_1) = \{x : \|x\| \leq \gamma_1\} \subseteq Q \subseteq B(\gamma_2)$. Then

$$\gamma_1 \|x\| \leq f(x) \leq \gamma_2 \|x\|.$$

Subgradient method: $f(x_k) - f^* \leq \frac{L \|x_0 - x^*\|}{\sqrt{k+1}}$.

We have: $L = \gamma_2$, $\|x_0\| \leq \|x^*\| \leq \frac{1}{\gamma_1} f^* \leq \frac{1}{\gamma_1} f(x_0)$.

Bounding the growth of $\tilde{\phi}$

Denote $l^{(i)} \stackrel{\text{def}}{=} \max\{\ln(i), \text{Out}(i)\}$, and $i \in \mathcal{N}$. Clearly,

$$\sum_{i=1}^n l_i \leq \sum_{i=1}^n [\ln(i) + \text{Out}(i)] = 2m.$$

For $T \in \mathcal{R}_S$ define the Euclidean metric

$$\|T\|_E = \left[\sum_{k=1}^q \sum_{i \neq k} l^{(i)} \left(\theta_{k,i}^* T_k^{(i)} \right)^2 \right]^{1/2}.$$

Bounding the growth of $\tilde{\phi}$

Denote $l^{(i)} \stackrel{\text{def}}{=} \max\{\ln(i), \text{Out}(i)\}$, and $i \in \mathcal{N}$. Clearly,

$$\sum_{i=1}^n l_i \leq \sum_{i=1}^n [\ln(i) + \text{Out}(i)] = 2m.$$

For $T \in \mathcal{R}_S$ define the Euclidean metric

$$\|T\|_E = \left[\sum_{k=1}^q \sum_{i \neq k} l^{(i)} \left(\theta_{k,i}^* T_k^{(i)} \right)^2 \right]^{1/2}.$$

Main Theorem: For any $T \in \mathcal{C}$ we have

$$\frac{1}{\sqrt{2mq}} \|T\|_E \leq \tilde{\phi}(T) \leq \sqrt{2mq} \|T\|_E.$$

Bounding the growth of $\tilde{\phi}$

Denote $l^{(i)} \stackrel{\text{def}}{=} \max\{\ln(i), \text{Out}(i)\}$, and $i \in \mathcal{N}$. Clearly,

$$\sum_{i=1}^n l_i \leq \sum_{i=1}^n [\ln(i) + \text{Out}(i)] = 2m.$$

For $T \in \mathcal{R}_S$ define the Euclidean metric

$$\|T\|_E = \left[\sum_{k=1}^q \sum_{i \neq k} l^{(i)} \left(\theta_{k,i}^* T_k^{(i)} \right)^2 \right]^{1/2}.$$

Main Theorem: For any $T \in \mathcal{C}$ we have

$$\frac{1}{\sqrt{2mq}} \|T\|_E \leq \tilde{\phi}(T) \leq \sqrt{2mq} \|T\|_E.$$

Consequences:

- The Lipschitz constant of function $\tilde{\phi}$ is bounded by $\sqrt{2mq}$.

Bounding the growth of $\tilde{\phi}$

Denote $l^{(i)} \stackrel{\text{def}}{=} \max\{\ln(i), \text{Out}(i)\}$, and $i \in \mathcal{N}$. Clearly,

$$\sum_{i=1}^n l_i \leq \sum_{i=1}^n [\ln(i) + \text{Out}(i)] = 2m.$$

For $T \in \mathcal{R}_S$ define the Euclidean metric

$$\|T\|_E = \left[\sum_{k=1}^q \sum_{i \neq k} l^{(i)} \left(\theta_{k,i}^* T_k^{(i)} \right)^2 \right]^{1/2}.$$

Main Theorem: For any $T \in \mathcal{C}$ we have

$$\frac{1}{\sqrt{2mq}} \|T\|_E \leq \tilde{\phi}(T) \leq \sqrt{2mq} \|T\|_E.$$

Consequences:

- The Lipschitz constant of function $\tilde{\phi}$ is bounded by $\sqrt{2mq}$.
- Optimal dual solution is bounded: $\frac{1}{\sqrt{2mq}} \|T^*\|_E \leq \tau^* \leq \tilde{\phi}(T)$.

Finding MCF by subgradient method $\mathcal{GM}(\delta)$

Denote $\mathcal{L} = \{T \in \mathcal{C} : \langle D, T \rangle = 1\}$.

Define projection $\pi_{\beta}(S) = \arg \min_{T \in \mathcal{L}} \{\langle S, T \rangle + \frac{\beta}{2} \|T\|_E^2\}$.

1. Set $k = 0$, $p = 0$, $N = 4m^2q^2 (2 + \frac{2.5}{\delta})^2$, $T_0 = \arg \min_{T \in \mathcal{L}} \|T\|_E$.

2. **Restart:** Set $S_k = 0$, $p := p + 1$, $\beta_p = \frac{\sqrt{N}}{\tilde{\phi}(T_k^*)}$, $\alpha_p = \frac{1}{2} \tilde{\phi}(T_k^*)$.

3. **Main cycle: for $i = 1$ to N do**

a) Compute $T_{k+1} = \pi_{\beta_p}(S_k)$, $\tilde{\phi}(T_{k+1})$, and $G_{k+1} \in \partial \tilde{\phi}(T_{k+1})$.

b) Set $S_{k+1} = S_k + G_{k+1}$, $k := k + 1$.

c) If $\tilde{\phi}(T_k^*) \leq \alpha_p$ then **go to** Step 2.

4. **Terminate**, if the last cycle was full.

In this scheme we denote $T_k^* = \arg \min_T \{\tilde{\phi}(T) : T = T_0, \dots, T_k\}$.

Complexity bounds for $\mathcal{GM}(\delta)$

Number of stages: $p \leq 2 + \log_2(qm)$.

Complexity of the stage: at most $4m^2q^2 \left(2 + \frac{2.5}{\delta}\right)^2$ iterations.

Complexity of iteration: $O(mq \ln q)$ operations.

Total: $O\left(\frac{1}{\delta^2} m^3 q^3 \ln q \ln m\right)$ operations.

Complexity bounds for $\mathcal{GM}(\delta)$

Number of stages: $p \leq 2 + \log_2(qm)$.

Complexity of the stage: at most $4m^2q^2 \left(2 + \frac{2.5}{\delta}\right)^2$ iterations.

Complexity of iteration: $O(mq \ln q)$ operations.

Total: $O\left(\frac{1}{\delta^2} m^3 q^3 \ln q \ln m\right)$ operations.

Thus, this is a Fully Polynomial-Time Approximation Scheme (FPTAS).

Smoothing technique

For $F = (f_1, \dots, f_k) \in R^{m \times q}$ define the Euclidean norm:

$$\|F\|_P^2 \stackrel{\text{def}}{=} \sum_{k=1}^q \sum_{j=1}^m \left(\frac{f_k^{(j)}}{\Theta_*^{(j,k)}} \right)^2.$$

Smoothing technique

For $F = (f_1, \dots, f_k) \in R^{m \times q}$ define the Euclidean norm:

$$\|F\|_P^2 \stackrel{\text{def}}{=} \sum_{k=1}^q \sum_{j=1}^m \left(\frac{f_k^{(j)}}{\Theta_{*}^{(j,k)}} \right)^2.$$

Let us introduce a linear operator $A(T)$, $T \in \mathcal{R}_S$, by identity

$$\langle A(T), F \rangle \equiv \sum_{k=1}^q \langle T_k, E f_k \rangle.$$

Recall: $\tilde{\phi}(T) = \max_F \{ \langle A(T), F \rangle : 0 \leq F \leq \Theta_*, F e \leq \bar{f} \}$.

For applying smoothing technique, we need an estimate for $\|A\|$.

Smoothing technique

For $F = (f_1, \dots, f_k) \in R^{m \times q}$ define the Euclidean norm:

$$\|F\|_P^2 \stackrel{\text{def}}{=} \sum_{k=1}^q \sum_{j=1}^m \left(\frac{f_k^{(j)}}{\Theta_*^{(j,k)}} \right)^2.$$

Let us introduce a linear operator $A(T)$, $T \in \mathcal{R}_S$, by identity

$$\langle A(T), F \rangle \equiv \sum_{k=1}^q \langle T_k, E f_k \rangle.$$

Recall: $\tilde{\phi}(T) = \max_F \{ \langle A(T), F \rangle : 0 \leq F \leq \Theta_*, F e \leq \bar{f} \}$.

For applying smoothing technique, we need an estimate for $\|A\|$.

Lemma

$$\|A\|_{E,P} \stackrel{\text{def}}{=} \max_{T,F} \{ \langle A(T), F \rangle : \|T\|_E \leq 1, \|F\|_P \leq 1 \} \leq 2.$$

Smoothing the dual objective

Let us define a smooth approximation of function $\tilde{\phi}$:

$$\tilde{\phi}_\mu(T) = \max_F \{ \langle A(T), F \rangle - \frac{\mu}{2} \|F\|_P^2 : 0 \leq F \leq \Theta_*, Fe \leq \bar{f} \},$$

where $\mu > 0$ is a smoothing parameter.

Smoothing the dual objective

Let us define a smooth approximation of function $\tilde{\phi}$:

$$\tilde{\phi}_\mu(T) = \max_F \{ \langle A(T), F \rangle - \frac{\mu}{2} \|F\|_P^2 : 0 \leq F \leq \Theta_*, Fe \leq \bar{f} \},$$

where $\mu > 0$ is a smoothing parameter.

Theorem

- $\tilde{\phi}_\mu$ has Lipschitz-continuous gradient with constant $L_\mu = \frac{4}{\mu}$.

Smoothing the dual objective

Let us define a smooth approximation of function $\tilde{\phi}$:

$$\tilde{\phi}_\mu(T) = \max_F \{ \langle A(T), F \rangle - \frac{\mu}{2} \|F\|_P^2 : 0 \leq F \leq \Theta_*, Fe \leq \bar{f} \},$$

where $\mu > 0$ is a smoothing parameter.

Theorem

- $\tilde{\phi}_\mu$ has Lipschitz-continuous gradient with constant $L_\mu = \frac{4}{\mu}$.
- The size of the primal feasible set is bounded by \sqrt{mq} .

Smoothing the dual objective

Let us define a smooth approximation of function $\tilde{\phi}$:

$$\tilde{\phi}_\mu(T) = \max_F \{ \langle A(T), F \rangle - \frac{\mu}{2} \|F\|_P^2 : 0 \leq F \leq \Theta_*, Fe \leq \bar{f} \},$$

where $\mu > 0$ is a smoothing parameter.

Theorem

- $\tilde{\phi}_\mu$ has Lipschitz-continuous gradient with constant $L_\mu = \frac{4}{\mu}$.
- The size of the primal feasible set is bounded by \sqrt{mq} .
- $\tilde{\phi}_\mu(T) \leq \tilde{\phi}(T) \leq \tilde{\phi}_\mu(T) + \frac{\mu}{2}mq$.

Fast gradient method

Define the *gradient mapping* (GM)

$$T_\mu(x) = \arg \min_{V \in \mathcal{L}} \left\{ \langle \nabla \tilde{\phi}_\mu(T), V - T \rangle + \frac{2}{\mu} \|V - T\|_P^2 \right\}.$$

Fast gradient method

Define the *gradient mapping* (GM)

$$T_\mu(x) = \arg \min_{V \in \mathcal{L}} \{ \langle \nabla \tilde{\phi}_\mu(T), V - T \rangle + \frac{2}{\mu} \|V - T\|_P^2 \}.$$

1. Set $k = 0$, $p = 0$, $N = 4mq \left(2 + \frac{2.5}{\delta}\right)$, $T_0 = \arg \min_{T \in \mathcal{L}} \|T\|_E$.
2. **Restart:** Set $S_k = 0$, $p := p + 1$, $\mu_p = \frac{8\tilde{\phi}(T_k^*)}{N+1}$, $\alpha_p = \frac{1}{2}\tilde{\phi}(T_k^*)$.
3. **Main cycle: for $i = 0$ to N do**
 - a) Compute $\nabla \tilde{\phi}_{\mu_p}(T_k)$ and $V_k = T_{\mu_p}(T_k)$.
 - b) Set $S_{k+1} = S_k + \frac{k+1}{2} \nabla \tilde{\phi}_{\mu_p}(T_k)$. Compute $Z_k = \pi_{4/\mu_p}(S_{k+1})$.
 - c) Set $T_{k+1} = \frac{2}{k+3} Z_k + \frac{k+1}{k+3} V_k$, $k := k + 1$.
 - d) If $\tilde{\phi}(T_k^*) \leq \alpha_p$ then **go to** Step 2.
4. **Terminate**, if the last cycle was full.

In this scheme $T_k^* = \arg \min_T \{ \tilde{\phi}(T) : T = T_0, \dots, T_k, V_0, \dots, V_k \}$.

Complexity bounds for Smoothing Technique

Number of stages: $p \leq 2 + \log_2(qm)$.

Complexity of the stage: at most $4mq \left(2 + \frac{2.5}{\delta}\right)$ iterations.

Complexity of iteration: $O(mq \ln q)$ operations

Complexity bounds for Smoothing Technique

Number of stages: $p \leq 2 + \log_2(qm)$.

Complexity of the stage: at most $4mq \left(2 + \frac{2.5}{\delta}\right)$ iterations.

Complexity of iteration: $O(mq \ln q)$ operations.

Total: $O\left(\frac{1}{\delta} m^2 q^2 \ln q \ln m\right)$ operations.

Auxiliary problems

Auxiliary problems

- Projection $\pi_\beta(S)$ or GM: $\min_{T \geq 0} \{ \langle S, T \rangle + \frac{\beta}{2} \|T\|_E^2 : \langle D, T \rangle = 1 \}$.

Auxiliary problems

- Projection $\pi_\beta(S)$ or GM: $\min_{T \geq 0} \{ \langle S, T \rangle + \frac{\beta}{2} \|T\|_E^2 : \langle D, T \rangle = 1 \}$.
- Function $\tilde{\phi}(T) = \max_F \{ \langle A(T), F \rangle : 0 \leq F \leq \Theta_*, Fe \leq \bar{f} \}$
 $= \sum_{j=1}^m \max_{u_j \in R^q} \{ \langle a_j(T), u_j \rangle : 0 \leq u_j \leq \Theta_*^T e_j, \langle e, u_j \rangle \leq \bar{f}^{(j)} \},$
where u_j are rows of F and a_j are some linear functions of T .

Auxiliary problems

- Projection $\pi_\beta(S)$ or GM: $\min_{T \geq 0} \{ \langle S, T \rangle + \frac{\beta}{2} \|T\|_E^2 : \langle D, T \rangle = 1 \}$.

- Function $\tilde{\phi}(T) = \max_F \{ \langle A(T), F \rangle : 0 \leq F \leq \Theta_*, Fe \leq \bar{f} \}$

$$= \sum_{j=1}^m \max_{u_j \in R^q} \{ \langle a_j(T), u_j \rangle : 0 \leq u_j \leq \Theta_*^T e_j, \langle e, u_j \rangle \leq \bar{f}^{(j)} \},$$

where u_j are rows of F and a_j are some linear functions of T .

- Smooth approximation

$$\tilde{\phi}_\mu(T) = \max_F \{ \langle A(T), F \rangle - \frac{\mu}{2} \|F\|_P^2 : 0 \leq F \leq \Theta_*, Fe \leq \bar{f} \}$$

$$= \sum_{j=1}^m \max_{u_j \in R^q} \{ \langle a_j(T), u_j \rangle - \frac{\mu}{2} \|D^{-1}(\Theta_*^T e_j) u_j\|_{(2)}^2 :$$

$$0 \leq u_j \leq \Theta_*^T e_j, \langle e, u_j \rangle \leq \bar{f}^{(j)} \}.$$

Auxiliary problems

- Projection $\pi_\beta(S)$ or GM: $\min_{T \geq 0} \{ \langle S, T \rangle + \frac{\beta}{2} \|T\|_E^2 : \langle D, T \rangle = 1 \}$.

- Function $\tilde{\phi}(T) = \max_F \{ \langle A(T), F \rangle : 0 \leq F \leq \Theta_*, Fe \leq \bar{f} \}$

$$= \sum_{j=1}^m \max_{u_j \in R^q} \{ \langle a_j(T), u_j \rangle : 0 \leq u_j \leq \Theta_*^T e_j, \langle e, u_j \rangle \leq \bar{f}^{(j)} \},$$

where u_j are rows of F and a_j are some linear functions of T .

- Smooth approximation

$$\tilde{\phi}_\mu(T) = \max_F \{ \langle A(T), F \rangle - \frac{\mu}{2} \|F\|_P^2 : 0 \leq F \leq \Theta_*, Fe \leq \bar{f} \}$$

$$= \sum_{j=1}^m \max_{u_j \in R^q} \{ \langle a_j(T), u_j \rangle - \frac{\mu}{2} \|D^{-1}(\Theta_*^T e_j)u_j\|_{(2)}^2 :$$

$$0 \leq u_j \leq \Theta_*^T e_j, \langle e, u_j \rangle \leq \bar{f}^{(j)} \}.$$

Dual feasible sets: a box intersected by one linear constraint.

Auxiliary problems

- Projection $\pi_\beta(S)$ or GM: $\min_{T \geq 0} \{ \langle S, T \rangle + \frac{\beta}{2} \|T\|_E^2 : \langle D, T \rangle = 1 \}$.

- Function $\tilde{\phi}(T) = \max_F \{ \langle A(T), F \rangle : 0 \leq F \leq \Theta_*, Fe \leq \bar{f} \}$

$$= \sum_{j=1}^m \max_{u_j \in \mathbb{R}^q} \{ \langle a_j(T), u_j \rangle : 0 \leq u_j \leq \Theta_*^T e_j, \langle e, u_j \rangle \leq \bar{f}^{(j)} \},$$

where u_j are rows of F and a_j are some linear functions of T .

- Smooth approximation

$$\tilde{\phi}_\mu(T) = \max_F \{ \langle A(T), F \rangle - \frac{\mu}{2} \|F\|_P^2 : 0 \leq F \leq \Theta_*, Fe \leq \bar{f} \}$$

$$= \sum_{j=1}^m \max_{u_j \in \mathbb{R}^q} \{ \langle a_j(T), u_j \rangle - \frac{\mu}{2} \|D^{-1}(\Theta_*^T e_j) u_j\|_{(2)}^2 :$$

$$0 \leq u_j \leq \Theta_*^T e_j, \langle e, u_j \rangle \leq \bar{f}^{(j)} \}.$$

Dual feasible sets: a box intersected by one linear constraint.

Complexity: $m \times \ln m$ or $m \times q \ln q$.

Conclusion

Conclusion

- We developed new FPTAS for MCFP.

Conclusion

- We developed new FPTAS for MCFP.
- The best of these schemes has complexity $O^*(\frac{1}{\delta} m^2 q^2)$ operations. (Best known?)

Conclusion

- We developed new FPTAS for MCFP.
- The best of these schemes has complexity $O^*(\frac{1}{\delta} m^2 q^2)$ operations. (Best known?)
- In these schemes the auxiliary problems are trivial.

Conclusion

- We developed new FPTAS for MCFP.
- The best of these schemes has complexity $O^*(\frac{1}{\delta} m^2 q^2)$ operations. (Best known?)
- In these schemes the auxiliary problems are trivial.

Open questions:

- We developed new FPTAS for MCFP.
- The best of these schemes has complexity $O^*(\frac{1}{\delta} m^2 q^2)$ operations. (Best known?)
- In these schemes the auxiliary problems are trivial.

Open questions:

- Can we have better performance on the optimization stage? (Choice of the method, Euclidean metric, etc.)

- We developed new FPTAS for MCFP.
- The best of these schemes has complexity $O^*(\frac{1}{\delta} m^2 q^2)$ operations. (Best known?)
- In these schemes the auxiliary problems are trivial.

Open questions:

- Can we have better performance on the optimization stage? (Choice of the method, Euclidean metric, etc.)
- Can we have cheaper preliminary computations? We need only an estimate for the maximal flows.