Contents

CHAPTER 1

Multicommodity Flows

Many flow optimization problems we encounter in real life are more complex than the ordinary min cost flow problem we studied in the last chapter. In particular, quite often we have to deal with a number of "overlapping" flows simultaneously. For example, in an electronic network, there are several data streams to be routed through the network simultaneously. Similarly, in a traffic network, we usually have a number of source-sink or *origin-destination pairs* (O-D *pairs*) and traffic flows between them. Thus let us assume that, instead of a single source-sink pair (r, s), we are given a finite number of O-D pairs (r_k, s_k), $k = 1, \dots, K$. For each O-D pair there is a corresponding *demand* $d_k \geq 0$ of flow from r_k to s_k. For example, in an electronic network, d_k would specify the amount of data to be routed from r_k to s_k. Correspondingly, we introduce flow variables $\mathbf{x}^{(k)} \in \mathbb{R}^E$, $k = 1, \dots, K$, with each $\mathbf{x}^{(k)}$ being a flow value d_k from r_k tp s_k. The goods (data, traffic) to be routed between the various O-D pairs are treated as different *commodities*. So we also refer to $\mathbf{x}^{(k)}$ as a *flow of commodity* k, $k = 1, \dots, K$. This explains the term *multicommodity flows*. As in the ordinary (single commodity) case, each edge $e \in E$ has an associated cost for sending one unit of flow through e. These edge costs may depend on the commodity. So let us assume that we are given edge costs $\mathbf{c}^{(k)} \in \mathbb{R}^E$, $k = 1, \dots, K$.

The flows $\mathbf{x}^{(k)}$ give rise to a *total flow* $\mathbf{x} = \sum_k \mathbf{x}^{(k)}$ in the network. Assuming capacities $\mathbf{u} \in \mathbb{R}^E$, the total flow is subject to capacity constraints $\mathbf{x} \leq \mathbf{u}$. (This is what causes the additional difficulty. If there were separate capacities $\mathbf{u}^{(k)}$ for each commodity k, our problem would decompose into K independent min cost flow problems.) These capacity constraints $\sum_k \mathbf{x}^{(k)} \leq \mathbf{u}$ are also referred to as *bundle constraints*.

Summarizing, the min cost problem for multicommodity flows can be stated as

$$
\begin{aligned}
z^* &= \min \sum_k \mathbf{c}^{(k)} \mathbf{x}^{(k)} \\
& \quad A\mathbf{x}^{(k)} = \mathbf{b}^{(k)} \\
& \quad \sum_k \mathbf{x}^{(k)} \leq \mathbf{u} \qquad k = 1, \dots, K \\
& \quad \mathbf{x}^{(k)} \geq \mathbf{0},
\end{aligned}
\tag{1.1}
$$

where $\mathbf{b}^{(k)} \in \mathbb{R}^V$ has components d_k in position s_k and $-d_k$ in position r_k and is zero elsewhere. (To simplify the notation, we write $\mathbf{c}^{(k)}\mathbf{x}^{(k)}$ instead of $(\mathbf{c}^{(k)})^T\mathbf{x}^{(k)}$.)

EX. 1.1. *Consider the network consisting of 3 nodes and 6 edges as in figure 1.1 below. There are 3 O-D pairs (r_k, s_k). All capacities are equal to 1 and edges costs are equal to 3 on the direct links (r_k, s_k) and 1 otherwise (for each commodity).*

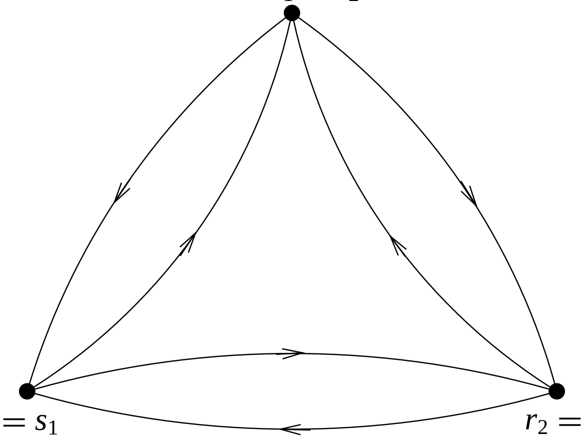

FIGURE 1.1. Three O-D pairs (r_k, s_k).

Solve the corresponding min cost multicommodity flow problem.

Of course, 1.1 could be solved directly as an LP. In practice, however, at least for large instances with many O-D pairs, this approach is computationally infeasible. Alternative solution methods have been proposed, taking advantage of the special (network) structure of the constraints. We present some of them in the following.

Remark. In contrast to the single commodity case, the existence of *integral* optimum flows $\mathbf{x}^{(k)}$ in 1.1 is no longer guaranteed by the integrality of the problem data $\mathbf{u}$ and d_k (*cf.* Ex 1.1). Indeed, computing integral solution of (1.1) turns out to be NP-hard.

1.1. Capacity Allocation

Assume we split the total available capacity $\mathbf{u} \in \mathbb{R}^E$ into K capacity vectors $\mathbf{r}^{(k)} \geq \mathbf{0}$, one for each commodity k:

$$\mathbf{u} = \mathbf{r}^{(1)} + \ldots + \mathbf{r}^{(k)}. \tag{1.2}$$

Allocating capacity $\mathbf{r}^{(k)} \in \mathbb{R}^E$ to commodity k decomposes the corresponding min cost problem into K independent min cost flow problems

$$
z_k\left(\mathbf{r}^{(k)}\right) \quad := \quad \begin{array}{lcl} \min \ \mathbf{c}^{(k)}\mathbf{x}^{(k)} & & \\ A\mathbf{x}^{(k)} & = & \mathbf{b}^{(k)} \\ \mathbf{0} \le \mathbf{x}^{(k)} & \le & \mathbf{r}^{(k)} \end{array} \tag{1.3}
$$

So a *capacity allocation* $\mathbf{r} = \left(\mathbf{r}^{(1)}, \ldots, \mathbf{r}^{(K)}\right)$ results in a corresponding total cost

$$
z(\mathbf{r}) = \sum_k z_k\left(\mathbf{r}^{(k)}\right). \tag{1.4}
$$

Problem (1.1) is now equivalent to the minimization of $z(\mathbf{r})$ subject to (1.2). The function $z(\mathbf{r})$ is a piecewise linear convex function (*cf.* Ex. 1.2), so we may apply the subgradient method (*cf.* Appendix) to solve the minimization problem. The subgradients of z can essentially be read off from the dual in (1.3), as explained below.

Ex. 1.2. *Show that* $z : (\mathbb{R}^E)^k \to \mathbb{R}$ *defined by (1.4) is a piecewise linear convex function.*

Subgradient Computation. In order to apply the subgradient method for minimizing $z(\mathbf{r})$, subject to (1.2), we must be able to compute subgradients of z restricted to the affine space

$$
H := \{\mathbf{r} = \left(\mathbf{r}(1), \ldots, \mathbf{r}^{(K)}\right) \mid \mathbf{u} = \sum_k \mathbf{r}^{(k)}\} \subset (\mathbb{R}^E)^K.
$$

Given $\overline{\mathbf{r}} \in H$ with finite $z(\overline{\mathbf{r}}) \in \mathbb{R}$ (*i.e.*, assuming that all problems (1.3) are feasible for $\mathbf{r} = \overline{\mathbf{r}}$), we are thus to compute a vector $\overline{\mathbf{p}} \in (\mathbb{R}^E)^K$ such that

$$
z(\mathbf{r}) \ge z(\overline{\mathbf{r}}) + \overline{\mathbf{p}}^T(\mathbf{r} - \overline{\mathbf{r}}) \qquad \forall \mathbf{r} \in H.
$$

Such a subgradient $\overline{\mathbf{p}}$ is easily obtained from an optimum dual solution, as we will see. First note that, by LP duality,

$$
\begin{array}{rclcrcl}
z(\overline{\mathbf{r}}) = \min \displaystyle\sum_k \mathbf{c}^{(k)}\mathbf{x}^{(k)} & & & = & \max \displaystyle\sum_k \mathbf{b}^{(k)}\mathbf{y}^{(k)} - \overline{\mathbf{r}}^{(k)}\mathbf{w}^{(k)} & & \\
A\mathbf{x}^{(k)} & = & \mathbf{b}^k & & \mathbf{y}^{(k)}A - \mathbf{w}^{(k)} & \le & \mathbf{c}^{(k)} \\
\mathbf{0} \le \mathbf{x}^{(k)} & \le & \overline{\mathbf{r}}^{(k)} & & \mathbf{w}^{(k)} & \ge & \mathbf{0}
\end{array}
$$

If $\overline{\mathbf{y}} = (\overline{\mathbf{y}}^{(1)}, \ldots, \overline{\mathbf{y}}^{(K)})$ and $\overline{\mathbf{w}} = (\overline{\mathbf{w}}^{(1)}, \ldots, \overline{\mathbf{w}}^{(K)})$ are optimum dual solutions, we thus find (with $\mathbf{b} = (\mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(K)})$)

$$
z(\overline{\mathbf{r}}) = \mathbf{b}^T\overline{\mathbf{y}} - \overline{\mathbf{r}}^T\overline{\mathbf{w}}.
$$

Since $(\overline{\mathbf{y}}, \overline{\mathbf{w}})$ is dually feasible for any $\mathbf{r} \in H$, we see that

$$
z(\mathbf{r}) \ge \mathbf{b}^T\overline{\mathbf{y}} - \mathbf{r}^T\overline{\mathbf{w}}, \qquad \mathbf{r} \in (\mathbb{R}^E_+)^K.
$$

Hence

$$z(\mathbf{r}) \geq z(\overline{\mathbf{r}}) - \overline{\mathbf{w}}^T(\mathbf{r} - \overline{\mathbf{r}}), \qquad \mathbf{r} \in (\mathbb{R}_+^E)^K.$$

Showing that $-\overline{\mathbf{w}}$ is a subgradient of $z : (\mathbb{R}_+^E)^K \to \mathbb{R}$. We claim that a subgradient of $z_{|H}$ at $\overline{\mathbf{r}}$ can be obtained by projecting $\overline{\mathbf{w}}$ onto the subspace $H_0 \leq (\mathbb{R}^E)^K$ corresponding to H. Indeed, let $\overline{\mathbf{p}} \in H_0$ denote the projection of $-\overline{\mathbf{w}}$ on H_0, *i.e.*, there exists some $\overline{\mathbf{q}} \in (\mathbb{R}^E)^K$ such that $-\overline{\mathbf{w}} = \overline{\mathbf{p}} + \overline{\mathbf{q}}$ and $\overline{\mathbf{q}} \perp H_0$. Then for any $\mathbf{r} \in H$ we have $\mathbf{r} - \overline{\mathbf{r}} \in H_0$ and hence $\overline{\mathbf{q}}^T(\mathbf{r} - \overline{\mathbf{r}}) = 0$, *i.e.*, $-\overline{\mathbf{w}}^T(\mathbf{r} - \overline{\mathbf{r}}) = \overline{\mathbf{p}}^T(\mathbf{r} - \overline{\mathbf{r}})$. So, indeed

$$z(\mathbf{r}) \geq z(\overline{\mathbf{r}}) + \overline{\mathbf{p}}^T(\mathbf{r} - \overline{\mathbf{r}}), \qquad \mathbf{r} \in H,$$

as claimed.

Remark. Recall that the dual variables $\overline{\mathbf{w}}^{(k)}$ can be interpreted as marginal costs, *i.e.*, $\overline{\mathbf{w}}_e^{(k)}$ is the prize we would be willing to pay for increasing $\overline{\mathbf{r}}_e^{(k)}$ by one unit. The subgradient method proceeds by moving from a current $\overline{\mathbf{r}}$ into the direction of the negative subgradient to a new capacity allocation $\overline{\mathbf{r}}' = \overline{\mathbf{r}} + \delta\overline{\mathbf{p}}$ for some stepsize $\delta > 0$. So it seeks to move as close as possible (subject to $\overline{\mathbf{p}} \in H_0$) into direction $\overline{\mathbf{w}} = (\overline{\mathbf{w}}^{(1)}, \ldots, \overline{\mathbf{w}}^{(k)})$.

1.2. Dualizing the bundle constraints

Lagrangian relaxation provides another way of decomposing (1.1) into K independent subproblems. Dualizing the bundle constraints with Lagrangian multipliers $\mathbf{w} \in \mathbb{R}_+^E$, we arrive at the relaxation

$$\begin{array}{rrcl} (1.5) \qquad & L(\mathbf{w}) & := & \min \sum_k (\mathbf{c}^{(k)} + \mathbf{w}^T)\mathbf{x}^{(k)} - \mathbf{w}^T\mathbf{u} \\ & A\mathbf{x}^{(k)} & = & \mathbf{b}^{(k)} \\ & \mathbf{x}^{(k)} & \geq & \mathbf{0} \end{array}$$

For given $\mathbf{w} \geq \mathbf{0}$, problem (1.5) is equivalent to K independent min cost path problems, relative to the modified edge costs $\widetilde{\mathbf{c}}^{(k)} = \mathbf{c}^{(k)} + \mathbf{w}$. So $L(\mathbf{w})$ is easy to compute, given $\mathbf{w} \geq \mathbf{0}$. The function $L : \mathbb{R}_+^E \to \mathbb{R}$ is a piecewise linear concave function (*cf.* Ex 1.3). So, again, the optimum z^* of (1.1) can be found by solving a convex minimization (concave maximization) problem, the Lagrangian dual

$$(1.6) \qquad z^* = \max_{\mathbf{w} \geq \mathbf{0}} L(\mathbf{w}).$$

And, again, the subgradient technique is the right choice for solving (1.6).

Ex. 1.3. *Show that $L : \mathbb{R}_+^E \to \mathbb{R}$ is a piecewise linear concave function.*

Subgradient computation. Given $\overline{\mathbf{w}} \geq \mathbf{0}$, we compute corresponding optimal solutions $\overline{\mathbf{x}}^{(k)}$ of (1.5), so that

$$L(\overline{\mathbf{w}}) = \sum_k \mathbf{c}^{(k)}\overline{\mathbf{x}}^{(k)} + \overline{\mathbf{w}}^T(\sum_k \overline{\mathbf{x}}^{(k)} - \mathbf{u}).$$

For arbitrary $\mathbf{w} \geq \mathbf{0}$ we have, by definition of $L(\mathbf{w})$,

$$L(\mathbf{w}) \leq \sum_k \mathbf{c}^{(k)}\overline{\mathbf{x}}^{(k)} + \mathbf{w}^T(\sum_k \overline{\mathbf{x}}^{(k)} - \mathbf{u}).$$

Hence

$$L(\mathbf{w}) \leq L(\overline{\mathbf{w}}) + (\sum_k \overline{\mathbf{x}}^{(k)} - \mathbf{u})^T(\mathbf{w} - \overline{\mathbf{w}}),$$

revealing $\mathbf{d} = \sum\limits_k \overline{\mathbf{x}}^{(k)} - \mathbf{u}$ as a subgradient of the concave function L at $\overline{\mathbf{w}}$. The subgradient method for maximizing the concave function L would (essentially) move from the current $\overline{\mathbf{w}}$ into direction $\mathbf{d} = \sum \overline{\mathbf{x}}^{(k)} - \mathbf{u}$. Interpreting $w_e \geq 0$ as an additional edge cost (*toll*) on edge $e \in E$, the subgradient method would thus increase the current edge toll $\overline{w}_e \geq 0$ in case e is *overloaded, i.e.,* $\sum \overline{x}_e^{(k)} > u_e$ and decrease the current edge toll $\overline{w}_e > 0$ whenever the edge capacity is not completely used up by the current flows $\overline{\mathbf{x}}^{(k)}$.

EX. 1.4. *Analyze how the subgradient method proceeds to solve the problem from Ex. 1.1, starting with* $\mathbf{w}^{(0)} = \mathbf{0}$ *and step sizes* $\delta_t \to 0$, $\sum\limits_t \delta_t = \infty$.

1.3. Path Decomposition (Simplex Method with Column Generation Technique)

Let $\mathcal{P}_k$ denote the set of (incidence vectors of) directed $r_k - s_k$ paths. As we assume $\mathbf{c}^{(k)} \geq \mathbf{0}$, the flow $\mathbf{x}^{(k)}$ solving (1.1) can be assumed to be acyclic and hence decompose into simple path flows

$$\mathbf{x}^{(k)} = \sum_{P \in \mathcal{P}_k} \lambda_P P.$$

(Recall that we use P to denote both a path $P \subset E$ and its incidence vector in $\{0, \pm 1\}^E$.)

We may thus restate (1.1) as follows

$$(1.7)\qquad \begin{aligned} \min \sum_k \sum_{P\in\mathcal{P}_k} \lambda_P \mathbf{c}^{(k)}(P) & \\ \sum_k \sum_{\substack{P\in\mathcal{P}_k,\\ P\ni e}} \lambda_P \le u_e & \qquad , e\in E \\ \sum_{P\in\mathcal{P}_k} \lambda_P = d_k & \qquad , k=1,\ldots,K \\ \lambda_P \ge 0 & \end{aligned}$$

The problem formulation (1.7) is an LP with only $m+K$ constraints – as opposed to $m+n\,K$ constraints in (1.1). On the other hand, (1.8) involves a huge number of variables $\lambda_P \ge 0$. Correspondingly, the dual of (1.7) has a huge number of constraints, one for each path $P \in \cup\mathcal{P}$, and a small number $m+K$ of variables. Let us denote the dual variables corresponding to the demand constraints in (1.7) by γ_k, $k=1,\ldots,K$ and the dual variables corresponding to the capacity constraints by $w_e \ge 0$, $e\in E$. The dual of (1.7) then becomes

$$(1.8)\qquad \begin{aligned} \max \sum_k d_k\gamma_k - \mathbf{u}^T\mathbf{w} & \\ \gamma_k - \mathbf{w}(P) \le \mathbf{c}^{(k)}(P) & \qquad , P\in\mathcal{P}_k, k=1,\ldots,K. \\ \mathbf{w} \ge \mathbf{0} & \end{aligned}$$

Remark. Just like in 1.2, the $w_e \ge 0$ may be interpreted as edge tolls. The variables γ_k, as we shall see, represent min path costs, relative to the modified edge weights $\widetilde{\mathbf{c}}^{(k)} = \mathbf{c}^{(k)} + \mathbf{w}$.

The important observation to make is that we can check dual feasibility of $(\gamma, \mathbf{w})$ very easily by solving K min cost path problems relative to the modified edge costs $\widetilde{\mathbf{c}}^{(k)}$. This allows us to implement the Simplex Method for solving (1.7) rather efficiently in practice.

To work this out in detail, assume (λ_P) is a current feasible basic solution of (1.7). Then (λ_P) satisfies some of the capacity constraints, say, those corresponding to $e_1,\ldots,e_t$, $t\ge 0$, with equality. Being a basic solution, λ then has at most $K+t$ nonzero components, say,

$$\begin{aligned} \lambda_P \ge 0 &\quad \text{for} \quad P\in\mathcal{B} = \cup\mathcal{B}_k \subseteq \cup\mathcal{P}_k, \\ \lambda_P = 0 &\quad \text{for} \quad P\notin\mathcal{B}, \end{aligned}$$

with $|\mathcal{B}| = K+t$.

We find a complementary dual solution $(\gamma, \mathbf{w})$ by setting $w_e = 0$ for $e\in E\backslash\{e_1,\ldots,e_t\}$ and computing the remaining $K+t$ dual variables $w_{e_1},\ldots,w_{e_t}$ and $\gamma_1,\ldots,\gamma_K$ so that

$$\gamma_k - \mathbf{w}(P) = \mathbf{c}^{(k)}(P), \qquad P \in \mathcal{B}_k, k = 1, \ldots, K.$$

In other words, all dual constraints corresponding to basic variables $\lambda_P \geq 0$, $P \in \mathcal{B}$, are satisfied with equality. Due to complementarity, the primal (λ_P) and the dual $(\gamma, \mathbf{w})$ then have the same objective value (*cf.* Ex 1.5). So in case $(\gamma, \mathbf{w})$ is dually feasible, we have arrived at a primal-dual pair of optimal solutions.

To check dual feasibility, it suffices to check the dual path constraints

$$\gamma_k - \mathbf{w}(P) \leq \mathbf{c}^{(k)}(P), \qquad P \in \mathcal{P}_k.$$

Indeed, assume that $(\gamma, \mathbf{w})$ satisfies all these path constraints, but some w_e are negative. Then increasing these negative w_e to 0 would result in a dually feasible solution $(\gamma, \mathbf{w}')$ with an even larger objective value, which is impossible.

Hence, in case $(\gamma, \mathbf{w})$ is not dually feasible, we can easily detect a violated path constraint

$$\gamma_k - \mathbf{w}(P) > \mathbf{c}^{(k)}(P)$$

for some $P \in \mathcal{P}_k$ by min cost path computations. We then add P to the current basis $\mathcal{B}$ and proceed to a new basis $\mathcal{B}' = \mathcal{B} + P$ (in case the corresponding basis solution (λ'_P) satisfies $t+1$ capacity constraints with equality) or to $\mathcal{B}' = \mathcal{B} + P - P'$, where the path P' that is to leave the current basis is determined in the usual way.

Remark. The constraint matrix of (1.8) essentially has columns corresponding to the paths $P \in \mathcal{P}$. The advantage of the Simplex approach described above is due to the fact that a Simplex step can be carried out without maintaining the (huge) matrix explicitely. Rather, in each step we only maintain the current basis $\mathcal{B}$ and generate a new path P (a new column) to enter the basis. For this reason, the above approach is an instance of *Simplex Method with Column Generation*.

Ex. 1.5. *Show that the primal* (λ_P) *and the corresponding complementary dual* $(\gamma, \mathbf{w})$ *yield the same objective value.*