'MATH+ECON+CODE' MASTERCLASS ON COMPETITIVE EQUILIBRIUM: WALRASIAN EQUILIBRIUM WITH SUBSTITUTES

Alfred Galichon (NYU)

Spring 2018

Day 6, May 26 2018: network problems with congestion and capacity constraints

- Congestion externalities
- Wardrop equilibrium
- Braess' paradox
- Min-cut, max flow theorem
- Ford-Fulkerson algorithm

- Koopmans (1949). "Optimum utilization of the transportation system." *Econometrica*.
- Braess (1968) "Uber ein paradoxon der verkehrsplanung". *Unternehmensforschung*.
- Dial (1971) "A probabilistic multipath traffic assignment model which obviates path enumeration". *Transportation Research*.
- Sheffi and Powell (1982) (1982). "An Algorithm for the Traffic Assignment Problem with Random Link Costs". *Networks*.
- Wardrop (1952) "Some theoretical aspects of road traffic research". *Proc. Inst. Civ. Eng.*
- Beckmann, McGuire, and Winsten (1956). *Studies in Economics of Transportation*. Yale University Press.
- Koutsoupias and Papadimitriou (1999). "Worst-case equilibria". *16th Annual Symposium on Theoretical Aspects of Computer Science*.
- Roughgarden and Tardos (2002) "How bad is selfish routing?". *Journal of the ACM*.

- Elias, Feinstein, and Shannon (1956). "A note on the maximum flow through a network", *IRE Transactions on Information Theory*.
- Ford and Fulkerson (1956). "Maximal flow through a network." *Canadian Journal of Mathematics*.
- Dantzig and Fulkerson (1956). "On the Max-Flow MinCut Theorem of Networks." *Ann. Math. Studies*.
- Dantzig and Fulkerson (1964). "On the max-flow min-cut theorem of networks." *RAND corporation*.

Section 1

Congestion externalities

- In the min-cost flow problem problem, we were minimizing a linear transportation cost $\mathcal{W}(\pi)$ under feasibility constraints, i.e.

$$\begin{aligned} &\min \mathcal{W}(\pi) \\ &s.t.\ \pi_{ij} \geq 0 \\ &\mathcal{N}\pi = b \end{aligned}$$

- We now would like to relax the assumption that our total total cost function $\mathcal{W}$ should be linear with respect to π. We shall take $\mathcal{W}$ as a separable function

$$\mathcal{W}(\pi) = \sum_{(i,j) \in A} K_{ij}(\pi_{ij})$$

where $K_{ij}(.)$ are real valued functions, one for each arc.
- This allows us to model *positive network spillovers*, which is the case where there are positive externalities, captured by the choice of $K_{ij}(x)$ as concave function, which means that path from i to j becomes less and less costly the more people go through it.
- Negative externalities, or *congestion effect*, are captured by a choice of convex function for $K_{ij}(x)$. Throughout the sequel, we shall assume that this is the case.

- Assume that $\mathcal{W}$ is a convex function. Then the primal value of the optimal transportation problem on the network

$$\begin{aligned} &\min \mathcal{W}(\pi) \\ &s.t.\ \pi \geq 0 \\ &\mathcal{N}\pi = b \end{aligned} \tag{1}$$

coincides with its dual value, which is

$$\max_{w} \sum_{i} w_i b_i - \mathcal{W}^* \left(w' \mathcal{N} \right) \tag{2}$$

where

$$\left(w' \mathcal{N} \right)_{ij} = w_j - w_i$$

and $\mathcal{W}^*$ is the convex conjugate function to $\mathcal{W}$, i.e.

$$\mathcal{W}^* (\kappa) = \sup_{\pi_{ij} \geq 0} \left(\sum_{(i,j) \in A} \pi_{ij} \kappa_{ij} - \mathcal{W}(\pi) \right). \tag{3}$$

- This follows from a min-max argument, as one has

$$
\begin{aligned}
& \min_{\pi \geq 0} \max_{w} \mathcal{W}(\pi) + w'(b - \mathcal{N}\pi) \\
& = \max_{w} w'b + \min_{\pi \geq 0} \mathcal{W}(\pi) - w'\mathcal{N}\pi \\
& = \max_{w} w'b - \max_{\pi \geq 0} w'\mathcal{N}\pi - \mathcal{W}(\pi) \\
& = \max_{w} w'b - \mathcal{W}^{*}\left(w'\mathcal{N}\right).
\end{aligned}
$$

- First, this problem is a generalization of the min-cost flow problem. Take
$$\mathcal{W}(\pi) = \sum_{(i,j)\in A} \pi_{ij} k_{ij}.$$
- Then, one has
$$\begin{aligned}\mathcal{W}^*(\kappa) &= 0 \text{ if } \kappa_{ij} \leq k_{ij} \text{ for all } (i,j) \in A \\ &= +\infty \text{ otherwise.}\end{aligned}$$

Hence, Equation (2) becomes

$$\max_w w'b$$
$$s.t.\ w'\mathcal{N} \leq k$$

recovering the min cost flow problem.

- We now give a more interesting important example. Consider the case where

$$\mathcal{W}(\pi) = \sum_{(i,j)\in A} \pi_{ij} k_{ij} + \sigma \sum_{(i,j)\in A} \pi_{ij} \ln \pi_{ij}.$$

- In that case, there is a vector $(w_i)_{i\in V}$ such that for each $(i,j) \in A$, the optimal flow π_{ij} satisfies the Schrödinger equation

$$\pi_{ij} = \exp\left(\frac{-k_{ij} + w_j - w_i - 1}{\sigma}\right), \tag{4}$$

where the w's exist, are unique up to an additive constant, and are a solution of

$$\max_w \sum_i w_i b_i - \sum_{(i,j)\in A} \sigma \exp\left(\frac{k_{ij} - w_j + w_i - \sigma}{\sigma}\right)$$

and the flow defined by Equation 4 is automatically feasible.

- The interpretation of this theorem is very interesting. The log-likelihood of a transition from i to j is proportional to minus the direct transportation cost $-k_{ij}$. Hence, all other things equal, all transitions are possible, but less costly transitions will be more likely than others. The potential w_i, on the other hand, adjusts π_{ij} so that it satisfies the feasibility constraint. Hence a terminal node with a high outgoing flow should "pump in" more mass, and therefore transitions to this node should receive higher probability.

- Proof: equation (3) becomes

$$\mathcal{W}^* (\kappa) = \sup_{\pi_{ij} \geq 0} \left(\sum_{(i,j) \in A} \pi_{ij} \left(\kappa_{ij} - k_{ij} - \sigma \ln \pi_{ij} \right) \right),$$

hence by first order conditions,

$$\kappa_{ij} - k_{ij} - \sigma \ln \pi_{ij} - \sigma = 0,$$

hence

$$\pi_{ij} = \exp \left(\frac{\kappa_{ij} - k_{ij} - \sigma}{\sigma} \right).$$

- Therefore

$$\mathcal{W}^* (\kappa) = \sum_{(i,j) \in A} \sigma \exp \left(\frac{\kappa_{ij} - k_{ij} - \sigma}{\sigma} \right)$$

and when $\kappa = w' \mathcal{N}$, one has $\kappa_{ij} = w_j - w_i$, thus

$$\pi_{ij} = \exp \left(\frac{w_j - w_i - k_{ij} - \sigma}{\sigma} \right),$$

The first order conditions associated to Equation (2), one gets

$$b_k = \frac{\partial \mathcal{W}^* \left(w' \mathcal{N} \right)}{\partial w_k}$$

thus

$$b_k = \sum_{a \in A} \frac{\partial \mathcal{W}^*}{\partial \kappa_a} \left(w' \mathcal{N} \right) \mathcal{N}_{ka},$$

hence

$$\begin{aligned} b_k = & \sum_{a \text{ arrives at } k} \exp\left(\frac{\kappa_a - k_a - \sigma}{\sigma} \right) \\ & - \sum_{a \text{ leaves from } k} \exp\left(\frac{\kappa_a - k_a - \sigma}{\sigma} \right) \end{aligned}$$

which is exactly the feasibility equation.

We now consider the individual decision problem, sometimes called "selfish routing problem". Consider the cost of adding transporting one incremental amount of mass δb in the network from source nodes S to terminal ones T. Let $\delta \pi$ the incremental flow generated.
Assume that the transportation cost of shipping $\delta \pi_{ij}$ through arc (i, j) is a function of the degree of saturation of the network: $k_{ij}\left(\pi_{ij}\right) \delta \pi_{ij}$, where $k_{ij}\left(.\right)$ are functions defined over each arcs and assumed to be increasing (in order to model congestion). Clearly, any incremental shipper will face a linear optimization cost with cost $k_{ij} = K'_{ij}\left(\pi_{ij}\right)$. This rules out cycles, and suboptimal paths in the network flow decomposition and this motivates the notion of a Wardrop equilibrium.

Definition. π is a Wardrop equilibrium if given any flow decomposition of π

$$\pi = \sum_{\rho \in \mathcal{P}} h_{\rho} 1 \{a \in \rho\} + \sum_{\mu \in \mathcal{C}} g_{\mu} 1 \{a \in \mu\},$$

then:
(i) $g_{\mu} = 0$ for all cycles μ, and
(ii) any path ρ with $h_{\rho} > 0$ from a source to a terminal node is optimal with respect to cost $k_{ij}(\pi_{ij})$.

π is a Wardrop equilibrium if and only if it solves problem (1)

$$\min_{\pi \geq 0} \sum_{ij} K_{ij}(\pi_{ij}) \qquad (5)$$
$$s.t.\ \mathcal{N}\pi = b$$

where K_{ij} is a primitive of k_{ij}, i.e. $K'_{ij}(x) = k_{ij}(x)$.
The first oder conditions of problem (5), coincide with those of

$$\min_{\hat{\pi} \geq 0} \sum_{ij} k_{ij} \hat{\pi}_{ij}$$
$$s.t.\ \mathcal{N}\hat{\pi} = b$$

where $k_{ij} = K'_{ij}(\pi_{ij})$. Thus Wardrop equilibria and optimizers of problem (1) coincide.

Note that π is not optimal. Indeed, the optimal π minimizes instead

$$\min_{\hat{\pi}\geq 0} \sum_{ij} k_{ij}\left(\hat{\pi}_{ij}\right)$$
$$s.t.\ \mathcal{N}\hat{\pi} = b$$

which is a different problem, unless the cost functions k_{ij} are linear.
The function

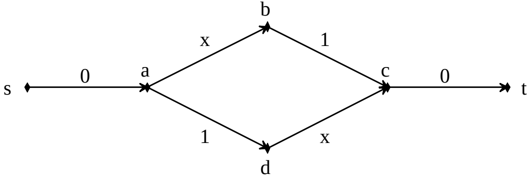
$$l_{ij}\left(x\right) = \frac{k_{ij}\left(x\right)}{x} = \frac{K'_{ij}\left(x\right)}{x}$$

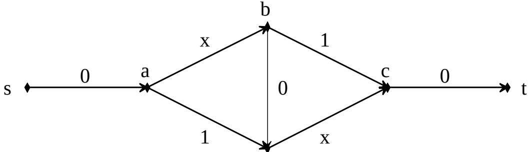
which captures the cost per unit of traffic is called the *latency function*.
With this definition, the optimal π minimizes

$$\min_{\hat{\pi}\geq 0} \sum_{ij} \hat{\pi}_{ij} l_{ij}\left(\hat{\pi}_{ij}\right) \tag{6}$$
$$s.t.\ \mathcal{N}\hat{\pi} = b.$$

which is clearly analoguous to (5), but l_{ij} is in general different from K_{ij}.
The loss of social welfare due to the difference between the optimal π and the equilibrium π is called in the literature the *price of anarchy* (Koutsoupias and Papadimitriou, 1999). It can be theoretically bounded.

Consider Figure 18, where the functions $k_{ij}(x)$ are indicated along the arcs. Thus there is no congestion effect in arcs (a, d) which costs one whatever the traffic is; and there is congestion effect in arcs (a, b) which cost π_{ab} when π_{ab} is the flow through that arc.





One would like to move one unit from node s to node t. In the first picture, the unique Wardrop equilibrium consists in splitting the flow into two halves, one on the path (s, a, b, c, t). Total cost per infinitesimal unit of mass is $3/2$ either way, hence total cost is $3/2$ and coincides with the optimum. Let us now consider the second picture, where one has simply added a free arc to the network from b to d. This obviously does not change the optimal flow, and one would anticipate that expanding possibilities has no reverse effect. It turns out that it actually *worsens* the situation. Indeed, irrespective of $x < 1$, the path (s, a, b, d, c, t) is now a shortest path, thus the only Wardrop equilibrium has now all traffic through that path – with a cost of 2.

Section 2

Capacity constraints

- In the max-flow problem, one defines a capacity $\bar{\mu}_a$ associated with each arc a in the network. Given the vector of outgoing flow s, a feasible flow is a vector $\mu \geq 0$ that should not only satisfy the mass balance equation $\nabla^\intercal \mu = s$, but also the capacity constraint $\mu_a \leq \bar{\mu}_a$ for all $a \in \mathcal{A}$.
- Assume w.l.o.g. that the total mass of source nodes (and hence the total mass of target nodes) is one, that is $\sum_{z:s_z>0} s_z = 1$. The max-flow problem is the problem of determining the highest $t \in \mathbb{R}$ such that there exists a feasible flow associated with ts. That is

$$\max_{t,\mu \geq 0} \{t\}$$
$$s.t.\ \nabla^\intercal \mu = ts$$
$$\mu \leq \bar{\mu}$$

- Consider $\Gamma \subseteq \mathcal{X} \times \mathcal{Y}$, and look for a perfect matching μ_{xy} along Γ between marginal distributions (n_x) and (m_y) such that $\sum_x n_x = \sum_y m_y$, i.e. such that

$$\sum_{y \in \mathcal{Y}} \mu_{xy} = n_x, \; \sum_{x \in \mathcal{X}} \mu_{xy} = m_y, \; \mu_{xy} > 0 \implies xy \in \Gamma.$$

- Create one origin node o and one destination node d, and set arcs ox such that $\mu_{ox} = n_x$, $\mu_{yd} = m_y$, and $s_z = 1\{z = d\} - 1\{z = 0\}$. Then it is easy to see that there is a perfect matching if and only if the maximum flow from o to d is one.

Theorem. The max-flow problem has dual expression

$$\begin{aligned} &\min_{p,\tau \geq 0} \bar{\mu}^\intercal \tau \\ &s.t.\ p^\intercal s = 1 \\ &\tau \geq \nabla p \end{aligned}$$

that is $\min_p \bar{\mu}^\intercal \left(\max \left\{ \nabla p, 0 \right\} : p^\intercal s = 1 \right)$.
Proof. Rewrite the max-flow problem as

$$\begin{aligned} &\max_{t,\mu \geq 0} \min_{p,\tau \geq 0} t + p^\intercal \nabla^\intercal \mu - t p^\intercal s + \bar{\mu}^\intercal \tau - \mu^\intercal \tau \\ &= \min_{p,\tau \geq 0} \bar{\mu}^\intercal \tau + \max_{t,\mu \geq 0} t \left(1 - p^\intercal s \right) + \mu^\intercal \left(\nabla p - \tau \right), QED. \end{aligned}$$

Now assume that there is only one source node z^o and one destination node z^d. Then $s = 1\{z = z^o\} - 1\{z = z^d\}$, so that the problem reformulates as

$$\min_p \sum_{xy \in \mathcal{A}} \bar{\mu}_{xy} \max\{p_y - p_x, 0\}$$
$$s.t.\ p_{z^o} = 0, p_{z^d} = 1.$$

The max-flow min-cut theorem expresses that one can take $p \in \{0, 1\}$, and so the problem becomes a min-cut problem.