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

Alfred Galichon (NYU)

Spring 2018
Day 1, May 21, 2018: the smooth case
Block 1. Equilibrium and optimization

- Schedule: Mon 5/21– Sat 5/26, 2018, 9am-1pm and 2pm-3pm.
- Location: Courant Institute (Warren Weaver Hall, 251 Mercer) WWH 202.
- Office hours: by apointment (my email: ag133@nyu.edu).
- Course webpage: http://alfredgalichon.com/mec_equil/
- Students (including auditors) need to register on Albert (MATH-GA 2840.002 or ECON-GA 3002.015)
- Texts (optional):
  - Ortega and Rheinboldt (1970). *Iterative Solution of Nonlinear Equations in Several Variables*. SIAM.
  - Galichon (2016). *Optimal Transport Methods in Economics*. Princeton.

- A. Galichon: professor of economics and of mathematics at NYU
- Octavia Ghelfi: NYU econ grad student (econ theory; focusing on matching markets and school choice problems)
- Keith O'Hara: NYU econ grad student (econometrics and coding; R and C++ guru)
- Yifei Sun: NYU math grad student (machine learning; joining Facebook research this summer)

- Jointly offered in the Math and Econ programs. Self-contained for both audiences
- Teaching format: 6 days, morning = theory; afternoon=coding presentations and assignments.
  - coding application related to the theory just seen
  - students are expected to write their own code to be pushed on a Github repository
- Programming: our demos will be done in R and the support will be in R only, but you are welcome to use the language of your choice e.g. Matlab, C++, Python, Julia...
- Questions?

- Provides the conceptual basis of competitive equilibrium with gross substitutes, along with various computational techniques (optimization problems, equilibrium problems). Shows how asynchronous parallel computation is adapted for the computation of equilibrium. Applications to hedonic equilibrium, multinomial choice with peer effects, and congested traffic equilibrium on networks.
- Describes analytical methods to analyze demand systems with gross substitutes (Galois connections, lattice programming, monotone comparative statics) and use them to study properties of competitive equilibrium with gross substitutes. Describe the Kelso-Crawford-Hatfield-Milgrom algorithm. Application to stable matchings, and equilibrium models of taxation.
- Derives models of bundled demand and analyze them using notions of discrete convexity and polymatroids. Application to combinatorial auctions and bundled choice.
- It will introduce tools from economic theory, mathematics, econometrics and computing, on a needs basis, without any particular prerequisite other than the equivalent of a first year graduate sequence in econ or in applied math.

- Day 1: the smooth case
- Day 2: lattice, isotonicity and gross substitutes
- Day 3: equilibrium transport
- Day 4: empirical models of matching
- Day 5: one-to-many matching models
- Day 6: traffic equilibrium with congestion

- (https://www.r-project.org/), Rstudio (https://www.rstudio.com/), and, for Windows users only, Rtools (https://cran.r-project.org/bin/windows/Rtools/).
- Jupyter (http://jupyter.org/), the R kernel for Jupyter (follow instructions from https://www.datacamp.com/community/blog/jupyter-notebook-r)
- gurobi (www.gurobi.com; you'll need to obtain an academic license, which is free if you're located on an educational domain).
- github (github.com), please make sure you have an account as you will need to push coding assignements in a repository.

On a sheet of paper, please indicate:

1. Your name and email
2. Your program, department and institution
3. Whether you are a registered student (RS) or an approved auditor (AA)
4. Whether you have taken the January 'm+e+c' course
5. What you are looking for in this course (2 sentences max).
6. From a scale of 1 to 5, whether you are less familiar (1) or more familiar (5) with the concepts below:

1. Linear programming
2. Legendre-Fenchel transforms
3. Min-cost flow problem
4. Strong set order
5. Backward induction
6. M-matrices
7. Envelope theorem
8. Complementary slackness
9. Gale and Shapley's algorithm
10. Tarski's fixed point theorem
11. Topkis' theorem
12. Gradient descent
13. Newton descent
14. Coordinate descent

Note: even if you anwer mostly "1", don't worry! These concepts will be explained in due course.

- Each afternoon will be dedicated to coding presentations followed by coding assignments. At the end of the day, wishing to be assessed are requested to push their code in a github directory.
- The course will be assessed on a pass/fail basis.

Block 1: Equilibrium and optimization

Learning objectives: block 1

- Understand the interplay between equilibrium and convex optimization
- Gradient descent algorithm
- Crash course on convex analysis

- Gul and Stacchetti (1999). "Walrasian equilibrium with gross substitutes". *Journal of Economic Theory*.
- Ekeland, Heckman, and Nesheim (2004). "Identification and Estimation of Hedonic Models," *Journal of Political Economy*.
- Heckman, Matzkin, and Nesheim (2010). "Nonparametric identification and estimation of nonadditive hedonic models," *Econometrica*.
- Dupuy, A., Galichon, A., and Henry, M. (2014). Entropy Methods for Identifying Hedonic Models. With Arnaud Dupuy and Marc Henry. *Mathematics and Financial Economics*, special issue in celebration of Ivar Ekeland's 70th birthday.

Section 1

Motivation

- The following model is described in Dupuy, G and Henry (2014). Consider a set of consumers considering the purchase on a housing good.
- Each consumer $i \in \mathcal{I}$ is represented by observable characteristics x_i (income, family size...), and each good $k \in \mathcal{K}$ is represented by observable characteristics z_k (amenities, surface,...). There are n_x consumers of type $x \in \mathcal{X}$ and l_z nondefault goods of type $z \in \tilde{\mathcal{Z}}$. In addition, there is a 0 "default" good – not buying, and we denote $\mathcal{Z}_0 = \tilde{\mathcal{Z}} \cup \{0\}$ the full set of goods.
- The equilibrium price of good z is p_z. Consumer i's utility if purchasing good z is

$$u_{xz} - p_z + \varepsilon_{iz}$$

and ε_{i0} if no good is purchased, where $(\varepsilon_{iz})_{z \in \mathcal{Z}_0} \sim$ i.i.d. Gumbel.

- The demand for good z is

$$D_z(p) = \sum_{x \in \mathcal{X}} n_x \frac{\exp(u_{xz} - p_z)}{1 + \sum_{z' \in \mathcal{Z}} \exp(u_{xz'} - p_{z'})}$$

- In the partial equilibrium, one takes the supply, i.e. the distribution of the goods $(l_z)_{z \in \mathcal{Z}_0}$ as given, and thus the equilibrium prices are determined by

$$D(p) = l.$$

- Consider promoters $j \in \mathcal{J}$ such that the observable type of j is $y_j \in \mathcal{Y}$ (institutional characteristics). There are m_y producers of type y. The cost for a producer j of type y to produce quality z is

$$p_z - c_{yz} + \eta_{jz}$$

and η_{0j} if no good is purchased, where $\left(\eta_{jz}\right)_{z \in \mathcal{Z}_0} \sim$ i.i.d. Gumbel.

- The supply for good z is

$$S_z\left(p\right) = \sum_{y \in \mathcal{Y}} m_y \frac{\exp\left(p_z - c_{yz}\right)}{1 + \sum_{z' \in \mathcal{Z}} \exp\left(p_{z'} - c_{yz'}\right)}$$

- In the general equilibrium, prices adjust demand and supply, and the equations are thus $D(p) = S(p)$, that is

$$E(p) = 0 \tag{1}$$

where $E : \mathbb{R}^{\mathcal{Z}} \to \mathbb{R}^{\mathcal{Z}}$ defined by $E(p) = S(p) - D(p)$ is the excess supply function.

- Once p has been determined by solving (1), one then computes the endogenous quality distribution $l = D(p) = S(p)$.
- This course is about how to compute equations of the type (1) in a large number of settings.

Section 2

Equilibrium and optimization

- A first way to solve for equilbrium is to recognize that in this context, the equilibrium equations $E(p) = 0$ can be interpreted as the first order conditions associated with an optimization problem.
- Indeed, $E(p) = \nabla W(p)$, where $W(p)$ is the indirect utility function

$$W(p) = W^D(p) + W^S(p), \text{ where}$$
$$\begin{cases} W^D(p) = \sum_{x \in \mathcal{X}} n_x \log\left(1 + \sum_{z \in \mathcal{Z}} \exp(u_{xz} - p_z)\right) \\ W^S(p) = \sum_{y \in \mathcal{Y}} m_y \log\left(1 + \sum_{z \in \mathcal{Z}} \exp(p_z - c_{yz})\right) \end{cases}$$

  which is smooth and convex, so the equilibrium equations can be interpreted as a convex optimization problem

$$\min_{p \in \mathbb{R}^{\mathcal{Z}}} W(p).$$

- Note that $D(p) = -\nabla W^D(p)$ and $S(p) = \nabla W^S(p)$.

- In the case of the partial equilibrium problem (exogenous supply), equilibrium is expressed as

$$\min_{p \in \mathbb{R}^{\mathcal{Z}}} W^D(p) + \sum_{z \in \mathcal{Z}} l_z p_z$$

i.e. one replaces $W^S(p)$ by $\sum_{z \in \mathcal{Z}} l_z p_z$.

- In partial equilibrium, we need to solve $D(p) = I$ where $D(p) = -\nabla W^D(p)$, hence

$$\nabla W^D(p) = -I.$$

- This can be formally inverted into as

$$p = \left(\nabla W^D\right)^{-1}(-I).$$

- The Jacobian of $\left(\nabla W^D\right)^{-1}$ is the inverse of the Jacobian of ∇W^D, thus it is the inverse of the Hessia of W^D; hence it is symmetric positive definite. Hence we should expect that $\left(\nabla W^D\right)^{-1}$ should be the gradient of a convex function. What it this convex function? we shall see that it is the Legendre-Fenchel transform of W^D.

- Recall the equilibrium equations in partial equilibrium

$$\nabla W^D(p) = -l. \tag{2}$$

- Note that equation (2) can be seen as the first order conditions to an optimization problem, namely, if we define the *Legendre-Fenchel transform* of W^D as

$$\left(W^D\right)^*(h) = \max_{p \in \mathbb{R}^{\mathcal{Z}}} \left\{ \sum_z h_z p_z - W^D(p) \right\}$$

  we get by the envelope theorem, that

$$\nabla \left(W^D\right)^* = \left(\nabla W^D\right)^{-1}$$

- Therefore, the equilibrium prices that solves equilibrium equations (2) are given by

$$p = \nabla \left(W^D\right)^* (-l).$$

- In our starting example, we had

$$\begin{cases} W^D(p) = \sum_{x \in \mathcal{X}} n_x \log\left(1 + \sum_{z \in \mathcal{Z}} \exp\left(u_{xz} - p_z\right)\right) \\ W^S(p) = \sum_{y \in \mathcal{Y}} m_y \log\left(1 + \sum_{z \in \mathcal{Z}} \exp\left(p_z - c_{yz}\right)\right) \end{cases}$$

- One can show (exercise) that

$$\left(W^D\right)^*(h) = \max_{\mu_{xz} \geq 0} \left\{ \sum_{xz} \mu_{xz} u_{xz} - \sum_{xz} \mu_{xz} \ln \mu_{xz} \right\}$$
$$s.t. \sum_{z \in \mathcal{Z}} \mu_{xz} \leq n_x, \quad \sum_{x \in \mathcal{X}} \mu_{xz} = -h_z.$$

and

$$\left(W^S\right)^*(h) = \max_{\mu_{yz} \geq 0} \left\{ -\sum_{yz} \mu_{yz} c_{yz} - \sum_{yz} \mu_{yz} \ln \mu_{yz} \right\}$$
$$s.t. \sum_{z \in \mathcal{Z}} \mu_{yz} \leq m_y, \quad \sum_{y \in \mathcal{Y}} \mu_{yz} = h_z.$$

- General equilibrium expresses as

$$D(p) = S(p) \tag{3}$$

- Calling l this common quantity, this is equivalent to

$$D^{-1}(l) = S^{-1}(l), \tag{4}$$

where equations (3) are in the price space, while equations (4) are in the quantity space.

- In our setting, we recall that $D(p) = -\nabla W^D(p)$ and $S(p) = \nabla W^S(p)$, so the equilibrium in the quantity space expresses as

$$\left(\nabla W^D\right)^{-1}(-l) = \left(\nabla W^S\right)^{-1}(l).$$

- As a result, I is determined by an optimization problem in the quantity space
$$\min_I \left\{ \left(W^D\right)^* (-I) + \left(W^S\right)^* (I) \right\}$$
which is a convex optimization problem whose first order conditions are exactly
$$D^{-1}(I) = S^{-1}(I).$$

- Note that in our starting example, the problem in the quantity space is

$$\max_{\substack{\mu_{xz}\geq 0\\ \mu_{yz}\geq 0}} \left\{ \sum_{xz} \mu_{xz} u_{xz} - \sum_{yz} \mu_{yz} c_{yz} - \sum_{xz} \mu_{xz} \ln \mu_{xz} - \sum_{yz} \mu_{yz} \ln \mu_{yz} \right\}$$
$$s.t. \sum_{z\in\mathcal{Z}} \mu_{xz} \leq n_x, \ \sum_{z\in\mathcal{Z}} \mu_{yz} \leq m_y, \ \sum_{x\in\mathcal{X}} \mu_{xz} = \sum_{y\in\mathcal{Y}} \mu_{yz}.$$

- Therefore, the problem in the quantity space is an optimal allocation problem.

Section 3

Algorithms

- We need to compute the optimization problem

$$\min_{p \in \mathbb{R}^{\mathcal{Z}}} W(p),$$

where $W(p) = W^D(p) + W^S(p)$, is a strictly convex smooth function.

- The first obvious method to compute this is gradient descent, a.k.a. tâtonnement:

$$p^{t+1} = p^t - \epsilon \nabla W\left(p^t\right)$$

where $\nabla W(p^t) = -D(p^t) + S(p^t) = E(p^t)$ is the excess supply function, that is

$$p^{t+1} = p^t - \epsilon E\left(p^t\right)$$

which has the natural interpretation: raise prices of overdemanded goods, decrease prices of oversupplied ones.

- Convergence of gradient descent is based on the following idea. Assume the time step tends to zero, so that in the continuous limit

$$\frac{dp(t)}{dt} = -E(p(t)).$$

- Then, recall that $E(p) = \nabla W(p)$, so that

$$\frac{dW(p(t))}{dt} = -\nabla W(p(t))^{\intercal} \frac{dp(t)}{dt} = -|E(p(t))|^2$$

therefore the value of W will decrease until the gradient becomes zero.

- This informal argument can be made precise and one can show that gradient descent converges to the minimum of $W(p)$, which is the equilibrium.

- A second algorithm is called coordinate descent. It amounts to minimizing iteratively $W(p)$ with respect to each coordinate p_z.
- Note that this will not always converge. For instance, if

$$W(p) = \max(p_1 + 1, p_2 + 1, 0)$$

  and if the algorithm starts at $p = 0$, it may stay there forever yielding a value of 1, although the value of the minimum is 0.
- However, we shall see in the next block that due to the *gross substitutes* property, this technique will work in our setting.