Mathematical tools for equilibrium

CE 392C

OUTLINE

1. Moving to larger problems
2. Fixed point problems
3. Some mathematical definitions: convex, compact, closed, bounded, continuous...
4. How can we write user equilibrium as a fixed point?
5. Variational inequalities and forces

The “trial and error” method doesn’t work well for realistic-sized networks:

- The Chicago regional network has 12982 nodes, 39018 links, and over 3 million OD pairs
- The Philadelphia network has 13389 nodes, 40003 links, and over 2 million OD pairs
- The Austin network has 7388 nodes, 18961 links, and around 1 million OD pairs.

Further, the number of *paths* in these networks is much, much larger.

You do *not* want a trial-and-error method for these networks. Later in the class we’ll discuss methods which scale better.

We will take a detour into optimization and other mathematical techniques which help us formulate and solve traffic assignment on large networks. If your multivariable calculus is a bit rusty, I'd advise reviewing the following concepts (see Chapter 3 of the text):

- Dot products and their geometric interpretation
- First and second partial derivatives
- The gradient vector
- The Hessian matrix
- Multivariate chain rule

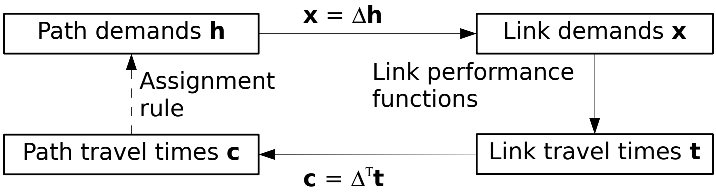
FIXED POINT PROBLEMS

There are three important questions you should be asking at this point:

- Does a user equilibrium solution always exist?
- If so, is the user equilibrium solution unique?
- Is there any practical way to find an equilibrium in large networks?

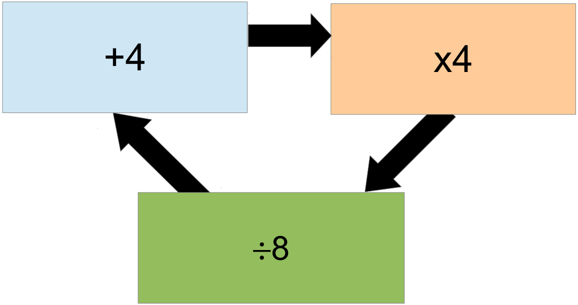
To answer these questions, we'll need some math. We will cover some basic results from fixed point problems, variational inequalities, and optimization.

In the last class, we interpreted user equilibrium as a “consistent” solution to this loop.



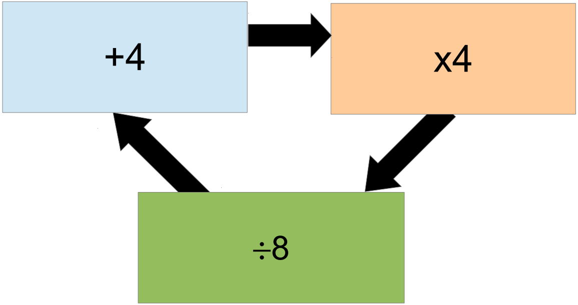
For example, if there was some function $R(\mathbf{c})$ which gives the path flows (route choice) as a function of path travel times, then we need to find $\mathbf{h}$ such that $\mathbf{h} = R(C(t(x(\mathbf{h}))))$

Consistency can be represented in a mathematical way:



Find a value which is unchanged when you go around the loop.

A clever person can solve this type of problem by denoting the values in the boxes as x, y, and z.



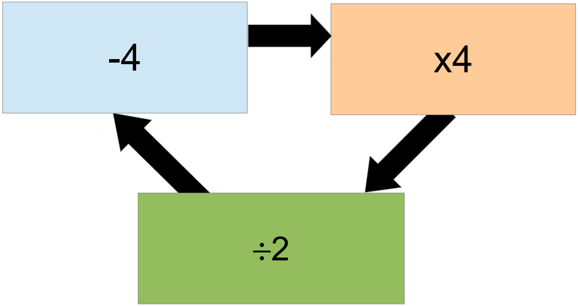
We then have $y = x + 4$, $z = 4y$, and $x = z/8$. Substituting them into each other, we have $x = \frac{1}{2}x + 2$ or $x = 4$.

A not-so-clever approach is to pick a starting value, then calculate through the loop iteratively.

$10 \to 7 \to 5.5 \to 4.75 \to 4.375 \to \ldots$ which converges to the correct answer (4).

Picking a different starting value: $1 \to 2.5 \to 3.25 \to 3.625 \to \ldots$ also converges to the same answer.

However, this simple approach won't always work.



Choosing the same starting value, we have $10 \to 12 \to 16 \to 24 \to 40 \to \ldots$ which diverges to $+\infty$.

The clever approach still works: $y = x - 4$, $z = 4y$, and $x = z/2$, so $x = 2x - 8$ and $x = 8$. If we used this as our starting value, the simple approach would work, but for any other value it will diverge.

This is an example of a fixed point problem. The more general definition is given below:

Consider some set X and a function f whose domain is X and whose range is contained in X. A **fixed point** of f is a value $x \in X$ such that $x = f(x)$.

(Note: it is **critical** that the function's range be contained in its domain. A fixed point problem does not make sense unless this is true.)

Fixed point theorems give us conditions on X and f which guarantee that a fixed point exists — for us, this will tell us when we known an equilibrium solution exists.

Brouwer's Theorem

If X is a compact convex set and f is a continuous function, then f has at least one fixed point.

This theorem is a bit frustrating in that it does give us any clue as to how to find this fixed point! But it must exist somewhere.

Mathematical definitions...

A set is **compact** if it is closed and bounded.

A set is **closed** if it contains all of its boundary points.

A set is **bounded** if it can be contained by a sufficiently large ball.

A set is **convex** if the line connecting any two points in the set lies within the set as well ($x \in X$ and $y \in X$ imply $\lambda x + (1 - \lambda)y \in X$ for all $\lambda \in [0, 1]$)

A function is **continuous** if at all points $y \in X$, $\lim_{x \to y} f(x)$ exists and is equal to $f(y)$.

To visualize the concept of fixed points, assume that $X = [0, 1]$.

A fixed point is anywhere $f(x)$ crosses the diagonal line $y = x$

One of the homework problems asks you to show that all of the conditions (closed, bounded, convex, continuous) are necessary for a fixed point to exist.

Application to traffic assignment

Does the traffic assignment problem satisfy the conditions of Brouwer's theorem?

Let H be the set of all feasible path flows. H is closed, bounded, and convex.

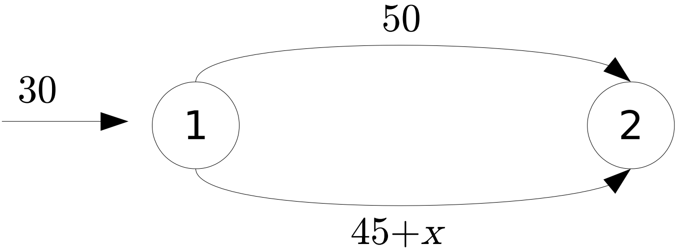
But what should $f : H \rightarrow H$ be? If paths are "tied" in travel time, then $R(C)$ can take infinitely many values.

If we stick with the fixed point approach, we can still make things work but we need to appeal to Kakutani's theorem instead.

Another approach, which is more useful for visualizing equilibrium problems, leads us to the variational inequality.

VARIATIONAL INEQUALITY

What does the set H look like in this example?



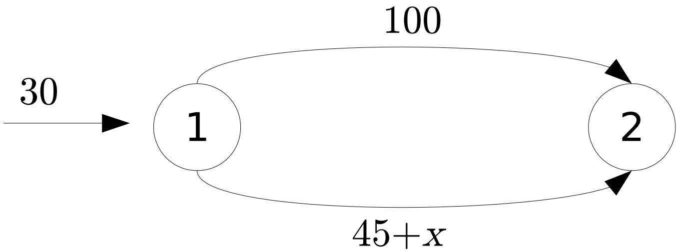
For each point $\mathbf{h}$ in H, associate a direction vector $-\mathbf{c}(\mathbf{h})$.

Think about this direction as a "force" which is pulling travelers (represented by the path flow vector) toward lower-cost alternatives (while staying within the feasible set).

Is there any point which is "stable" with respect to this force?

This force is not a physical force, and is primarily a device that will let us formulate the equilibrium problem. You can think of it is as a tendency for travelers to move toward lower-cost paths.

What if the problem is changed slightly?



What does H look like for larger problems?

In all of these cases, for a point $\hat{\mathbf{h}}$ to be stable, **the force $-\mathbf{c}(\hat{\mathbf{h}})$ cannot have any positive component in a direction which stays in H.**

Geometrically, this means that $-\mathbf{c}(\hat{\mathbf{h}})$ cannot make an acute angle with any vector $\mathbf{h} - \hat{\mathbf{h}}$ when $\mathbf{h} \in H$.

Equivalently, $-\mathbf{c}(\hat{\mathbf{h}}) \cdot (\mathbf{h} - \hat{\mathbf{h}}) \leq 0$ or $\mathbf{c}(\hat{\mathbf{h}}) \cdot (\hat{\mathbf{h}} - \mathbf{h}) \leq 0$ for all $\mathbf{h} \in H$.

This is a variational inequality: find $\hat{\mathbf{h}} \in H$ such that $\mathbf{c}(\hat{\mathbf{h}}) \cdot (\hat{\mathbf{h}} - \mathbf{h}) \leq 0$ for all $\mathbf{h} \in H$

The interpretation in terms of forces is intuitive; people want to move to lower cost paths, and an equilibrium is a “stable” point where nobody can reduce their travel times further.

Can we connect it with the principle of user equilibrium? That is, are the stable points of this system path flows where all used paths have equal and minimal travel times?

Theorem: $\hat{\mathbf{h}}$ solves the variational inequality $\mathbf{c}(\hat{\mathbf{h}}) \cdot (\hat{\mathbf{h}} - \mathbf{h}) \leq 0$ for all $\mathbf{h} \in H$ iff $\hat{\mathbf{h}}$ satisfies the principle of user equilibrium.

Part I: $\hat{\mathbf{h}}$ solves the VI if $\hat{\mathbf{h}}$ is a user equilibrium.

The VI can be rewritten $\mathbf{c}(\hat{\mathbf{h}}) \cdot \hat{\mathbf{h}} \leq \mathbf{c}(\hat{\mathbf{h}}) \cdot \mathbf{h}$.

The left hand side is another way to write TSTT at $\hat{\mathbf{h}}$: $\mathbf{c}(\hat{\mathbf{h}}) \cdot \hat{\mathbf{h}}$

The right hand side would give the TSTT at $\mathbf{h}$ *if the travel times were constant at the values corresponding to* $\hat{\mathbf{h}}$.

Since $\hat{\mathbf{h}}$ is UE, the only paths with positive h^π values have the least travel time out of all paths for that OD pair. So if the travel times stay constant, no matter what $\mathbf{h}$ is, the TSTT cannot be lower at $\mathbf{h}$ than at $\hat{\mathbf{h}}$. Thus $\hat{\mathbf{h}}$ solves the VI.

Theorem: $\hat{\mathbf{h}}$ solves the variational inequality $\mathbf{c}(\hat{\mathbf{h}}) \cdot (\hat{\mathbf{h}} - \mathbf{h}) \leq 0$ for all $\mathbf{h} \in H$ iff $\hat{\mathbf{h}}$ satisfies the principle of user equilibrium.

Part II: $\hat{\mathbf{h}}$ solves the VI only if $\hat{\mathbf{h}}$ is a user equilibrium.

If $\hat{\mathbf{h}}$ solves the VI, then if the travel times were held fixed, any other path flows $\hat{\mathbf{h}}$ would result in an equal or greater TSTT. This can only happen if the only used paths have equal and minimal travel time.

Existence

We are now in a position to use Brouwer's theorem. H is closed, bounded, and convex. We can define $f : H \to H$ as follows:

Let $f(\mathbf{h})$ be the location of the point ($\mathbf{h}$ after acted on by the force $-\mathbf{c}(\mathbf{h})$ for a short period of time: $f(\mathbf{h}) = \mathrm{proj}_H(\mathbf{h} - \mathbf{c}(\mathbf{h}))$ where proj_H means projection onto the feasible set H. If $\mathbf{c}$ is continuous, so is $f(\mathbf{h})$.

If $\hat{\mathbf{h}}$ solves the VI, then the point is unmoved by this force and $\hat{\mathbf{h}} = f(\hat{\mathbf{h}})$, that is, it is a fixed point of f.

Brouwer's theorem guarantees at least one fixed point of f, which is a solution of the VI, and which therefore satisfies the principle of user equilibrium.

OPTIMIZATION TECHNIQUES

You've seen some simple optimization problems before, in calculus:

"You have 60 ft of fencing. How can you use this fencing to enclose the largest possible rectangular area?"

There are much larger optimization problems which are used in many fields:

"You have $5 million to spend on pavement maintenance this year. Which maintenance actions should you perform on what roadway segments to maximize pavement condition for travelers?"

Why are we going in this direction?

Researchers have developed efficient methods for solving large-scale optimization problems with tens of thousands and even millions of variables. If we can reframe traffic assignment on a network as an optimization problem, we can use these methods as well.

These slides give an introduction and some highlights from the art of optimization. We'll shortly show how traffic assignment can fit into this framework.

Every optimization problem has three components:

Objective function: What are you trying to maximize or minimize?

Decision variables: What aspects of the problem can you control, both directly and indirectly?

Constraints: What restrictions are there on the possible values of the decision variables?

Example: You have 60 ft of fencing. How can you use this fencing to enclose the largest possible rectangular area?

Example: You are a private toll road operator. What toll should you charge to maximize your profit?

Optimization problems are often written in the following form:

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \text{Constraint} 1 \\ & \vdots \\ & \text{Constraint} m \end{aligned}$$

where **x** is a vector containing all of the decision variables. We often use X to denote the *feasible region* for the problem, that is, the set of **x** values which satisfy all of the constraints.

Some functions are easier to optimize than others. Set and function convexity play a very important role in determining which optimization problems are easier to solve.

FUNCTION CONVEXITY

Consider a function $f(x)$ of one variable, whose domain X is a convex set.

Geometrically, this function is convex if it "lies below its secants"

Mathematically, f is convex if, for every $x_1 \in X$ and $x_2 \in X$, and for every $\lambda \in [0, 1]$, we have

$$f(\lambda x_2 + (1 - \lambda)x_1) \leq \lambda f(x_2) + (1 - \lambda)f(x_1)$$

Furthermore, if f is differentiable, a function is convex iff it "lies above its tangents":

$$f(x_2) \geq f(x_1) + f'(x_1)(x_2 - x_1)$$

Furthermore, if f is twice differentiable, a function is convex iff

$$f''(x) \geq 0$$

for all $x \in X$

A function is *strictly convex* if these inequalities can be made strict.

Examples

Which of the following functions are convex? Strictly convex? (Assume their domain is convex.)

1. $f(x) = x^2$
2. $f(x) = 3x$
3. $f(x) = \sin x$
4. $f(x) = |x|$

When f is a function of multiple variables, the convexity conditions involving first and second derivatives must change.

The analogue of the first derivative is the *gradient* vector
$\nabla f = \begin{bmatrix} \partial f / \partial x_1 & \partial f / \partial x_2 & \cdots & \partial f / \partial x_n \end{bmatrix}^T$

The analogue of the second derivative is the *Hessian* matrix

$$Hf = \begin{bmatrix} \partial^2 f / \partial x_1^2 & \partial^2 f / \partial x_1 \partial x_2 & \cdots & \partial^2 f / \partial x_1 x_n \\ \partial^2 f / \partial x_2 \partial x_1 & \partial^2 f / \partial x_2^2 & \cdots & \partial^2 f / \partial x_2 x_n \\ \vdots & \vdots & \ddots & \vdots \\ \partial^2 f / \partial x_n \partial x_1 & \partial^2 f / \partial x_n \partial x_2 & \cdots & \partial^2 f / \partial x_n^2 \end{bmatrix}$$

For twice-differentiable multidimensional functions, f is convex if any of these equivalent conditions are satisfied:

1. For all x_1 and x_2 in X,

$$f(\lambda x_2 + (1-\lambda)x_1) \leq \lambda f(x_2) + (1-\lambda)f(x_1)$$

2. For all x_1 and x_2 in X,

$$f(x_2) \geq f(x_1) + \nabla f(x_1)^T(x_2 - x_1)$$

3. For all x in X, $H(x)$ is positive semidefinite (that is, $y^T H(x) y \geq 0$ for all vectors y).

For twice-differentiable multidimensional functions, f is *strictly* convex if any of these conditions are satisfied:

1. For all x_1 and x_2 in X,

$$f(\lambda x_2 + (1-\lambda)x_1) < \lambda f(x_2) + (1-\lambda)f(x_1)$$

2. For all x_1 and x_2 in X,

$$f(x_2) > f(x_1) + \nabla f(x_1)^T(x_2 - x_1)$$

3. Hf is positive definite (that is, $y^T H(x) y > 0$ for all nonzero vectors y).

Checking whether a matrix is positive definite or positive semidefinite can be tedious. In this class, though, many of the Hessians we will see are diagonal matrices. In this case, it's much easier to check:

A diagonal matrix is positive semidefinite iff all of its diagonal entries are nonnegative; a diagonal matrix is positive definite iff all of its diagonal entries are strictly positive.

Example

Which of these functions are convex? strictly convex?

1. $5x_1^2 + 2x_2^2$
2. $6x^3 + 4y^2$
3. $2x_1^2 - 2x_2^2$

CONVEX OPTIMIZATION

If the objective is a convex function, and the feasible region is a convex set, the following hold true:

- If x is a local minimum, it is also a global minimum.
- *(Uniqueness.)* If f is strictly convex, the global minimum is unique.
- *(Connectedness.)* The set of global minima is a convex set.

From here on out, we will assume that the objective function is convex (but not necessarily strictly convex)

ONE-DIMENSIONAL OPTIMIZATION

If there are *no* constraints and a single variable, $\min f(x)$ is easy to solve if f is differentiable.

Set $f'(x) = 0$ and solve for x.

Since f is convex we don't have to check whether this point is a minimum or a maximum, local or global, etc.

What happens if there is a nonnegativity constraint: $\min f(x)$ such that $x \geq 0$

There are two possibilities:

- $x \geq 0$ and $f'(x) = 0$ (same as before, just checking that x is feasible)
- $x = 0$ and $f'(x) \geq 0$ (the optimal point is at $x = 0$)

We want to summarize these conditions with equations. We need both x and $f'(x)$ to be nonnegative, but *at least one of them must be zero*.

The following equations do this:

$$\begin{aligned} x &\geq 0 \\ f'(x) &\geq 0 \\ xf'(x) &= 0 \end{aligned}$$

These are called the *optimality conditions* for the problem. If we can find a value of x that satisfies all three conditions, it is the optimal solution.

Example

Minimize $f(x) = x^2 + 3x + 5$ such that $x \geq 0$

MULTIPLE NONNEGATIVE VARIABLES

What if there is more than one decision variable, all of which must be nonnegative?

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

Using the same logic as before, we can derive the following optimality conditions which must hold *for every decision variable* x_i:

$$\begin{aligned} x_i &\geq 0 \\ \frac{\partial f(x)}{\partial x_i} &\geq 0 \\ x_i \frac{\partial f(x)}{\partial x_i} &= 0 \end{aligned}$$

This can be compactly written as

$$\mathbf{0} \leq \mathbf{x} \perp \nabla f(\mathbf{x}) \geq \mathbf{0}$$

Example

Minimize $f(x_1, x_2) = x_1^2 + x_2^2 + 3x_1 - 3x_2 + 5$ such that $x_1, x_2 \geq 0$

LINEAR EQUALITY CONSTRAINTS

Now, assume we have an optimization problem with multiple variables, a linear equality constraint, but no nonnegativity constraint.

An example is

$$\begin{aligned} \min_{x_1, x_2} \quad & x_1^2 + x_2^2 \\ \text{s.t.} \quad & x_1 + x_2 = 5 \end{aligned}$$

The technique of Lagrange multipliers can solve this problem.

The Lagrangian function $\mathcal{L}$ includes the original objective function, and a term for each constraint.

$$\mathcal{L}(x_1, x_2, \kappa) = x_1^2 + x_2^2 + \kappa(5 - x_1 - x_2)$$

where the new variable κ is a *Lagrange multiplier*.

You can think of the new term in the Lagrangian function as a "penalty" for violating the constraint.

The optimal solution of the original problem is a *stationary point* of the Lagrangian, that is, a place where $\nabla\mathcal{L}$ is zero.

Example

$$\begin{aligned} \min_{x_1, x_2} \quad & x_1^2 + x_2^2 \\ \text{s.t.} \quad & x_1 + x_2 = 5 \end{aligned}$$

What if we have *both* linear equality constraints and nonnegativity constraints on each variable?

$$\begin{aligned}
\min_{x_1,\ldots,x_n} \quad & f(x_1,\ldots,x_n) \\
\text{s.t.} \quad & \textstyle\sum_{i=1}^n a_{1i}x_i = b_1 \\
& \textstyle\sum_{i=1}^n a_{2i}x_i = b_2 \\
& \vdots \\
& \textstyle\sum_{i=1}^n a_{mi}x_i = b_m \\
& x_1,\ldots,x_n \geq 0
\end{aligned}$$

The Lagrangian function contains a multiplier and term for each constraint

$$\mathcal{L}(x_1,\ldots,x_n,\kappa_1,\ldots,\kappa_m) = f(x_1,\ldots,x_n) + \kappa_1\left(b_1 - \sum_{j=1}^n a_{1j}x_j\right) + \kappa_2\left(b_2 - \sum_{j=1}^n a_{2j}x_j\right) + \cdots + \kappa_m\left(b_m - \sum_{j=1}^n a_{mj}x_j\right)$$

We combine the Lagrangian approach with the optimality conditions from before, to obtain

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x_i} &\geq 0 && \forall i \in \{1, \ldots, n\} \\
\frac{\partial \mathcal{L}}{\partial \kappa_j} &= 0 && \forall j \in \{1, \ldots, m\} \\
x_i &\geq 0 && \forall i \in \{1, \ldots, n\} \\
x_i \frac{\partial \mathcal{L}}{\partial x_i} &= 0 && \forall i \in \{1, \ldots, n\}
\end{aligned}
$$

as the optimality conditions.