Perron's theorem

There is a famous theorem by the german mathematician Oskar Perron going back to 1907. It is extremely usefull in many situations, and it has important applications in many fields, just to mention s few:

Economics, theoretical physics, statistics. A more fancy application is Google's algortihm for ranging web-pages, which is based on the Perron-theorem.

The Perron theorem is about matrices with strictly positive entries and it states that the eigenvalue of such a matrix with maximal modulus (*a priori* it is a complex number) is real, and that its eigenspace is of dimesion one. Furthermore, there is an eigenvector whose entries all are strictly positive.

In these notes we treat a simplified version of Perron's, which we can prove without to much trouble by using the Banach fixed point theorem.

Our version is about what is called *probality matrices*. That is, matrices $P = (p_{ij})$ — of some size, say $n \times n$ — whose entries are real numbers between zero and one — *i.e.,* $0 \le p_{ij} \le 1$ — and whose column-sums all are equal to one — *i.e.,* $\sum_{i=1}^{n} p_{ij} = 1$.

The way of thinking about such a matrix is to regard it as a transition matrix for a system. Such a system consists of a *population* whose individuals can be in a certain *states*. The numbers between 1 and n form a numbering of the states, and the entries p_{ij} of the matrix are the probabilities that a member of the population being in state i swops to state j.

There are plenty of examples: For example, the population can be all the molecules in a fixed volume of gas, say hydrogen, in which case the states are the different excitation levels of the hydrogen molecule. A more mundane example: The population can be the set of TV-slaves in a given country and the states all possible TV-channels they have access to. And finally, we mention the Google example again: The population then being the population of the world having access to the internet, and the states are the set of web pages indexed by Google (which is quit a big number)!.

A vector $x = (x_1, \dots, x_n)$ — with $\sum_i x_i = 1$ — represents a distribution in percentage of the population among the states, and a vector satisfying $Px = x$, *i.e.,* an eigenvector with eigenvalue one, is a stable distribution; *i.e.,* one that does not change with time. And Perron says, that under certain conditions, such a stable distribution exists and is unique. In addition, the vector x has all components strictly positive; meaning that in the stable situation any state has some part of the population in it.

A conclusion like $Px = x$ immediatly sets our brain-vibrations in fixed-point-mode; and indeed, Perrons theorem follows rather easily from Banach's fixed point theorem. Which of course is the reason for these notes. One more comment about the applications: The iteration prosess in the proof of Banach's fixed point therem, also gives a good way to compute an approximation to the stabel eigenvector.

We shall mostly use what we call the Manhattan metric on $\mathbb{R}^n$. Being equivalent to the Euclidian metric, it does not change the topology: The two metrics have the same open and closed sets, the same continuous functions, the same convergent sequences *etc.*, but using it, the few computations we face, are much more agreeable.

A fixed point theorem

Allthough the Perron theorem folows from the Banach's fixed point theorem, we shall use a slightly different theorem (also given as an exercise in Tom's notes; 2.5 exercise 14) . Strengthening the hypothesis in Banach's theorem on the space allows a weakening of the hypothesis on the map:

Theorem 1 *Let X be a compact metric space with metric d and let $f\colon X \to X$ be a map. Assume that*

$$d(f(x), f(y)) < d(x, y) \tag{✣}$$

for all $x, y \in X$ with $x \neq y$. Then there is a unique point $x_0 \in X$ such that $f(x_0) = x_0$.

PROOF. The function $F(x) = d(x, f(x))$ is a continuous real valued function on X. Since X is compact, it achieves its minimum value at a point $x_0 \in X$. We claim that x_0 is a fixed point for f. Indeed, if $x_0 \neq f(x_0)$, the hypothesis (✣) above gives $F(f(x_0)) = d(f(x_0), f(f(x_0))) < d(x_0, f(x_0)) = F(x_0)$ which is absurd since $F(x_0) = d(x_0, f(x_0))$ was the minimal value of $F(x)$. Hence $x_0 = f(x_0)$.

Uniqueness follows since $x = f(x)$ and $y = f(y)$, we get $d(x, y) = d(f(x), f(y)) < d(x, y)$ unless $x = y$. □

Perron's theorem.

Theorem 2 *Let $P = (p_{ij})$ be a probability matrix, i.e., a matrix with $0 \leq p_{ij} \leq 1$ and $\sum_i p_{ij} = 1$ for $j = 1, \dots, n$. Assume that all the entries of P are strictly positive, i.e., $p_{ij} > 0$ for $1 \leq i, j \leq n$. Then there is an eigenvector for P with eigenvalue 1 all whose components are strictly positive. That is, there is an $x \in \mathbb{R}^n$ with $Px = x$ satisfying $x_i > 0$ for $1 \leq i \leq n$.*

Furthermore, if we impose $\sum_i x_i = 1$, the vector x is unique.

As we said, we are going to apply the above version of the Banach fixed point theorem, and for that we need a compact metric space. The one we are going to use, is what is called the *unit simplex* in $\mathbb{R}^n$, that is the set

$$\Delta = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid 0 \leq x_i \text{ for } 1 \leq i \leq n \text{ and } \sum_i x_i = 1\}. \qquad (✤)$$

We will give Δ the metric d which is the restriction of what popularly is called the Manhattan metric or the Taxi Cab metric on $\mathbb{R}^n$, and which is given by:

$$d(x, y) = \sum_i |x_i - y_i|.$$

It is easy to see that this metric is equivalent to the standard, Euclidian metric.

Lemma 1 *Δ is compact.*

Proof. The conditions $x_i \geq 0$ all define closed sets as well as the condition $\sum_i x_i = 1$, so Δ is closed in $\mathbb{R}^n$. Since $x_i \geq 0$, we see that each x_j satisfies $0 \leq x_j \leq \sum_i x_i = 1$ Hence Δ is also bounded, hence compact.

The Manhattan metric being equivalent to the Euclidian one, implies that Δ is also compact with respect to that metric. □

The next observation is:

Lemma 2 *If $x \in \Delta$, then $Px \in \Delta$; i.e., P defines a mapping $P \colon \Delta \to \Delta$.*

This is crucial since fixed point theorems deal with mappings from a set to itself.

Proof. We compute, using that the i-th coordinate of Px is $\sum_j p_{ij} x_j$:

$$\sum_i (\sum_j p_{ij} x_j) = \sum_j (\sum_i p_{ij} x_j) = \sum_j (\sum_i p_{ij}) x_j = \sum_j x_j = 1,$$

where the main trick is to change the order of summation. □

The next, and most important step, is to see that the mapping P satisfies the condition (✣) of the fixed point theorem above. We start with the following lemma, where we introduce the condition $\sum_j z_j = 0$; the reason is that it is satisfied by what interests us, namely the difference $z = x - y$ of two elements x and y from Δ (the sum of the components of both being one).

Lemma 3 *Let $z \in \mathbb{R}^n$ satisfy $\sum_j z_j = 0$, then*

$$\sum_i |\sum_j p_{ij} z_j| < \sum_i |z_i| \qquad (✥)$$

unless $z = 0$.

PROOF. We compute, using the triangle inequality and changing the order of summation:

$$\sum_i |\sum_j p_{ij}z_j| \leq \sum_i(\sum_j p_{ij}|z_j|) = \sum_j(\sum_i p_{ij})|z_j| = \sum_j |z_j|,$$

and are left to argue that there is strict inequality. Indeed if equality holds, then $\sum_i |\sum_j p_{ij}z_j| = \sum_i(\sum_j p_{ij}|z_j|)$ and hence $|\sum_j p_{ij}z_j| = \sum_j p_{ij}|z_j|$ since $|\sum_j p_{ij}z_j| \leq \sum_j p_{ij}|z_j|$.

But if the absolute value of a sum of some real numbers is equal to the sum of their absolute values, then those numbers all have the same sign. That means that the $p_{ij}z_j$'s — and hence the z_j's, as the p_{ij} are positive — are all of the same sign. But, since the z_i's add up to zero, that is impossible unless they are all equal to zero. Hence the strict inequality (✣) is established. □

We get immediatly

Lemma 4 *Any two elements $x, y \in \Delta$ with $x \neq y$ satisfy*

$$d(Px, Py) < d(x, y).$$

PROOF. Use lemma 3 with $z = x - y$; then $\sum_j z_j = \sum_j x_j - \sum_j y_j = 1 - 1 = 0$. □

Now we we can apply the fixed point theorem 1, and we get:

Theorem 3 *There is unique fixed point $x \in \Delta$ for the mapping P; i.e., a point with $Px = x$. All the coordinates of x are strictly positive, i.e., if $x = (x_1, \dots, x_n)$ then $x_i > 0$.*

PROOF. That there is a fixed point and that it is unique, follows from theorem 1: The set Δ is compact by lemma 1, and by lemma 4 above, the hypothesis (✣) is satisfied.

The only thing left, is the claim that $x_i > 0$ for i between 1 and n. The set Δ includes points not satisfying this, so we need to argue for it: Pick an i with $1 \leq i \leq n$. Now $x_i = \sum_j p_{ij}x_j$, and since all the $p_{ij} > 0$ and all the $x_j \geq 0$, it follows that either is $x_i > 0$ or all the x_j's are zero. The latter can not be, since $\sum_j x_j = 1$. □

Versjon: Monday, February 6, 2012 1:00:29 PM