VII.2 : Nonretraction and fixed point theorems

(**H**, §2.B; **M**, §55)

Recall that a continuous mapping $f : X \to Y$ is a **retract** if there is a continuous mapping $g : Y \to X$ such that $g \circ f$ is the identity on X, and a continuous mapping $p : X \to Y$ is a **retraction** if there is a continuous mapping $q : Y \to X$ such that $p \circ q$ is the identity on Y. It follows that the mappings q and f are 1–1 and the mappings p and g are onto; also, q is a retract and g is a retraction. Many but not all subspace inclusion mappings are retracts, and the following result shows the special natures of retracts.

PROPOSITION 1. *Suppose that the mapping $f : X \to Y$ is a retract. Then the induced maps in homology f_* are injections onto direct summands.*

Proof. The composite $g_* \circ f_* = (g \circ f)_*$ is the identity on $H_*(X)$. Therefore the identity $g_* f_*(a) = a$ implies that the kernel of f_* is zero so that f_* is 1–1. One can then check directly that each group $H_q(Y)$ is isomorphic to the direct sum of the image of f_* and the kernel of g_*.■

COROLLARY 2. *For every $k \geq 2$, the sphere S^{k-1} is not a retract of the disk D^k.*

This is true because the homology groups of D^k vanish in all positive dimensions but the group $H_{k-1}(S^{k-1})$ is nonzero.■

The Brouwer Fixed Point Theorem

At this point it is almost traditional to state and prove the Brouwer Fixed Point Theorem.

THEOREM 3. (Brouwer Fixed Point Theorem) *For all $n \geq 0$ every continuous map $f : D^n \to D^n$ has a* fixed point*; in other words, there is a point $\mathbf{x}$ in D^n such that $f(\mathbf{x}) = \mathbf{x}$.*

Proof. If $n = 1$ this is a fairly simple exercise in point set topology. Suppose that $i : S^0 \to D^1$ is the inclusion mapping and $r : D^1 \to S^0$ is such that $r \circ i$ is the identity on S^0. As noted before, it follows that r is onto; since D^1 is connected but S^0 is not, this is impossible and consequently there cannot be a continuous mapping r such that $r \circ i$ is the identity.

In the remaining cases, the standard proof is analogous to the argument on page 32 of Hatcher for the case $n = 2$. Assume that there is a continuous mapping $f : D^n \to D^n$ with no fixed point, so that $f(x) \neq x$ for all x. For each point $x \in D^n$, let $r(x) \in S^{n-1}$ be the unique point where the ray from $f(x)$ through x meets the boundary sphere. By construction $r(x) = x$ if $x \in S^{n-1}$, and if r is continuous then it follows that S^{n-1} is a retract of D^n because $r \circ i$ is the identity.

The argument in Hatcher states that the "continuity of r is clear"; although this seems reasonable, it is still necessary to check the continuity of the geometrically described mapping r retraction explicitly. For the sake of completeness we have written up the details in the Appendix to this section.

An application to matrices with nonnegative entries

The Brouwer Fixed Point Theorem and its generalizations play important roles in many branches of mathematics and their applications to other subjects. Here is one online reference for further information:

`http://en.wikipedia.org/wiki/Brouwer_fixed-point_theorem`

In these notes we shall only use the theorem to prove a result on eigenvalues and eigenvectors of matrices. The first step is the following elementary fact:

LEMMA 4. *Let X be a topological space which is homeomorphic to D^n for some $n \geq 0$. Then every continuous map $f : X \to X$ has a fixed point.*

Proof. Let $f : X \to X$ be continuous, and let $h : X \to D^n$ be a homeomorphism. Then $h \circ f \circ h^{-1}$ is a continuous map from D^n to itself and thus has a fixed point $\mathbf{p}$ by Brouwer's Theorem. In other words we have $h \circ f \circ h^{-1}(\mathbf{p}) = \mathbf{p}$. If we take $\mathbf{q} = h(\mathbf{p})$, straightforward computation shows that $f(\mathbf{q}) = \mathbf{q}$.■

THEOREM 5. (Perron – Frobenius) *Let $n > 1$, and let A be an $n \times n$ matrix which is invertible and has nonnegative entries. Then A has a positive eigenvalue λ such that λ has a nonzero eigenvector with nonnegative entries.*

Proof. Recall that the 1-norm on $\mathbf{R}^n$ is defined by $|\mathbf{x}|_1 = \sum_j |x_j|$, where the coordinates of $\mathbf{x}$ are given by $x_1, \cdots, x_n$. For each $\mathbf{x} \in \Delta_n$ (the standard simplex whose vertices are the unit vectors), define

$$f(\mathbf{x}) \quad = \quad (\,|A\mathbf{x}|_1)^{-1} \cdot A\mathbf{x} \;.$$

Observe that the coordinates of $A\mathbf{x}$ are all nonnegative because the entries of A and the coordinates of $\mathbf{x}$ are nonnegative, this vector is nonzero because A is invertible, and if $\mathbf{y}$ is a nonzero vector with nonnegative entries then $|\mathbf{y}|_1^{-1}\mathbf{y}$ must lie in Δ_n. Therefore we indeed have a continuous map f from the simplex to itself.

By the lemma, we know that f has a fixed point; in other words, there is some $\mathbf{v} \in \Delta_n$ such that

$$\mathbf{v} \quad = \quad (\,|A\mathbf{v}|_1)^{-1} \cdot A\mathbf{v}$$

and since the latter is equivalent to saying that $A\mathbf{v}$ is a positive multiple of $\mathbf{v}$, this completes the proof.■

COROLLARY 6. *In the setting of the theorem above, if all the entries of the matrix A are positive, then the eigenvector has positive entries.*

Proof. Let $\mathbf{y}$ be the eigenvector obtained in the theorem. Since $A\mathbf{y}$ is a positive scalar multiple of $\mathbf{y}$, it will suffice to prove that the entries of $A\mathbf{y}$ are all positive. But these entries are given by expressions of the form

$$z_i \quad = \quad \sum_j \; a_{i,j} y_j$$

and if we choose k such that $y_k \neq 0$ then it follows that $z_i \geq a_{i,k} y_k$; the right hand side is a product of

APPENDIX: The retraction in Brouwer's Theorem

As indicated before, the idea is relatively simple. We start with two distinct points $\mathbf{x}$ and $\mathbf{y}$ on the disk D^n and consider the ray starting with $\mathbf{y}$ and passing through $\mathbf{x}$; algebraically, this is the set of all points expressible as $\mathbf{y} + (1-t)\mathbf{x}$, where $t \geq 0$. Simple pictures strongly suggest that there is a unique scalar $t \geq 1$ such that $\mathbf{y} + (1-t)\mathbf{x}$ lies on S^{n-1}, if $\mathbf{x} \in S^{n-1}$ then $t = 1$ so that the point is equal to $\mathbf{x}$, and in fact the value of t is a continuous function of $(\mathbf{x},\mathbf{y})$. Our purpose here is to justify these assertions.

PROPOSITION. *There is a continuous function* $\rho : D^n \times D^n - \text{Diagonal} \to S^{n-1}$ *such that* $\rho(\mathbf{x},\mathbf{y}) = \mathbf{x}$ *if* $\mathbf{x} \in S^{n-1}$.

If we have the mapping ρ and f is a continuous map from D^n to itself without fixed points, then the retraction from D^n onto S^{n-1} is given by $\rho(\mathbf{x}, f(\mathbf{x}))$.

Proof of the proposition. It follows immediately that the intersection points of the line joining $\mathbf{y}$ to $\mathbf{x}$ are give by the values of t which are roots of the equation

$$|\mathbf{y} \ + \ t(\mathbf{x}-\mathbf{y})|^2 \quad = \quad 1$$

and the desired points on the ray are given by the roots for which $t > 1$. We need to show that there is always a unique root satisfying this condition, and that this root depends continuously on $\mathbf{x}$ andn $\mathbf{y}$.

We can rewrite the displayed equation as

$$|\mathbf{x}-\mathbf{y}|^2 t^2 \ + \ 2\langle \mathbf{y}, \mathbf{x}-\mathbf{y}\rangle t \ + \ (|\mathbf{y}|^2 - 1) \quad = \quad 0 \ .$$

If try to solve this nontrivial quadratic equation for t using the quadratic formula, then we obtain the following:

$$t \quad = \quad \frac{-\langle \mathbf{y}, \mathbf{x}-\mathbf{y}\rangle \ \pm \ \sqrt{\langle \mathbf{y}, \mathbf{x}-\mathbf{y}\rangle^2 \ + \ |\mathbf{x}-\mathbf{y}|^2 \cdot (1 - |\mathbf{y}|^2)}}{|\mathbf{x}-\mathbf{y}|^2}$$

One could try to analyze these roots by brute force, but it will be more pleasant to take a more qualitative viewpoint.

(a) *There are always two distinct real roots.* We need to show that the expression inside the square root sign is always a positive real number. Since $|\mathbf{y}| \leq 1$, the expression is clearly nonnegative, so we need only eliminate the possibility that it might be zero. If this happens, then each summand must be zero, and since $|\mathbf{y}-\mathbf{x}| > 0$ it follows that we must have both $\langle \mathbf{y}, \mathbf{x}-\mathbf{y}\rangle = 0$ and $1 - |\mathbf{y}|^2 = 0$. The second of these implies $|\mathbf{y}| = 1$, and the first then implies

$$\langle \mathbf{y}, \mathbf{x}\rangle \quad = \quad |\mathbf{y}|^2 \quad = \quad 1 \ .$$

If we combine this with the Cauchy-Schwarz Inequality and the basic condition $|\mathbf{x}| \leq 1$, we see that $|\mathbf{x}|$ must equal 1 and $\mathbf{x}$ must be a positive multiple of $\mathbf{y}$; these in turn imply that $\mathbf{x} = \mathbf{y}$, which contradicts our hypothesis that $\mathbf{x} \neq \mathbf{y}$. Thus the expression inside the radical sign is positive and hence there are two distinct real roots.

(b) *There are no roots t such that* $0 < t < 1$. The Triangle Inequality implies that

$$|\mathbf{y} \ + \ t(\mathbf{x}-\mathbf{y})| \quad = \quad |(1-t)\mathbf{y} \ + \ t\mathbf{x}| \quad \leq \quad (1-t)|\mathbf{y}| \ + \ t|\mathbf{x}| \quad \leq 1$$

so the value of the quadratic function

$$q(t) \quad = \quad |\mathbf{x}-\mathbf{y}|^2 t^2 \ + \ 2\langle \mathbf{y}, \mathbf{x}-\mathbf{y}\rangle t \ + \ (|\mathbf{y}|^2 - 1)$$

lies in $[-1, 0]$ if $0 < t < 1$. Suppose that the value is zero for some t_0 of this type. Since there are two distinct roots for the associated quadratic polynomial, it follows that the latter does not take a maximum value at t_0, and hence there is some t_1 such that $0 < t_1 < 1$ and the value of the function at t_1 is positive. This contradicts our observation about the behavior of the function, and therefore our hypothesis about the existence of a root like t_0 must be false.

(*d*) *There is one root of $q(t)$ such that $t \leq 0$ and a second root such that $t \geq 1$.* We know that $q(0) \leq 0$ and that the limit of $q(t)$ as $t \to -\infty$ is equal to $+\infty$. By continuity there must be some $t_1 \leq 0$ such that $q(t_1) = 0$. Similarly, we know that $q(1) \leq 0$ and that the limit of $q(t)$ as $t \to +\infty$ is equal to $+\infty$, so again by continuity there must be some $t_2 \geq 1$ such that $q(t_2) = 0$.

(*d*) *The unique root t satisfying $t \geq 1$ is a continuous function of $\mathbf{x}$ and $\mathbf{y}$.* This is true because the desired root is given by taking the positive sign in the expression obtained from the quadratic formula, and it is a routine algebraic exercise to check that this expression is a continuous function of $(\mathbf{x}, \mathbf{y})$.

(*e*) *If $|\mathbf{x}| = 1$, then $t = 1$.* This just follows because $|\mathbf{y} + 1(\mathbf{x} - \mathbf{y})| = 1$ in this case.

The proposition now follows by taking

$$\rho(\mathbf{x}, \mathbf{y}) \quad = \quad \mathbf{y} + t(\mathbf{x} - \mathbf{y})$$

where t is given as above by taking the positive sign in the quadratic formula. The final property shows that $\rho(\mathbf{x}, \mathbf{y}) = \mathbf{x}$ if $|\mathbf{x}| = 1$.■