The Contraction Mapping Lemma and the Inverse Function Theorem in Advanced Calculus

Lance D. Drager; Robert L. Foote

The American Mathematical Monthly, Vol. 93, No. 1 (Jan., 1986), 52-54.

(ii) P_k is parallel to P_0, $k = 1, \ldots, n$.
(iii) $P_k \cap I^n \neq \phi$, $k = 0, 1, \ldots, n$.
(iv) The distance of P_k from the origin is $k/\sqrt{n}$, $k = 0, 1, \ldots, n$.

Further, let S_{nk} be that portion of I^n bounded by P_{k-1} and P_k, and let $V(S_{nk})$ denote the volume of S_{nk}.

Then the following result can be proved.

THEOREM. $n!V(S_{nk}) = A(n,k)$.

Proof. Let $X_1, \ldots, X_n$ be n independent and identically distributed uniform random variables defined over $[0,1]$. Let

$$G(X_1, \ldots, X_n) = \frac{1}{n} \sum_{i=1}^{n} X_i = \overline{X}_n.$$

Then the probability density function of $\overline{X}_n$ takes the form

$$f_{\overline{X}_n}(x) = \frac{n^n}{(n-1)!} \sum_{j=0}^{n} (-1)^j \binom{n}{j} \left[c\left(x - \frac{j}{n} \right) \right]^{n-1}, \quad 0 \leqslant x \leqslant 1,$$

where $c(u) = 0$ if $u < 0$, and $c(u) = u$ if $u \geqslant 0$ (See [2], pp. 257–259 for a proof of this result.)

Elementary integration then yields

$$pr\left\{ \frac{(k-1)}{n} < \overline{X}_n < \frac{k}{n} \right\} = \int_{(k-1)/n}^{k/n} f_{\overline{X}_n}(x)\, dx$$
$$= \frac{1}{n!} \sum_{j=0}^{k} (-1)^j \binom{n+1}{j} (k-j)^n = A(n,k)/n!.$$

Finally, since $S_{nk} = \{(X_1, \ldots, X_n) | (k-1)/n < G(X_1, \ldots, X_n) < k/n\}$, the theorem follows.

Note that the result remains true (with some minor modifications) if we replace $(1, \ldots, 1)'$ by any vector parallel to a main diagonal of I^n.

References

1. L. Comtet, Advanced Combinatorics, Reidel, 1974.
2. M. G. Kendall and A. Stuart, The Advanced Theory of Statistics, 3rd ed., Hafner, 1969.

THE TEACHING OF MATHEMATICS

EDITED BY MARY R. WARDROP AND ROBERT F. WARDROP

Material for this department should be sent to Professor Robert F. Wardrop, Department of Mathematics, Central Michigan University, Mount Pleasant, MI 48859.

THE CONTRACTION MAPPING LEMMA AND THE INVERSE FUNCTION THEOREM IN ADVANCED CALCULUS

LANCE D. DRAGER AND ROBERT L. FOOTE
Department of Mathematics, Texas Tech University, Lubbock, TX 79409

In an undergraduate advanced calculus course, one of the important theorems a student sees is the inverse function theorem. The most elegant and general proof of this uses the contraction mapping lemma (see [1]). Unfortunately, time does not usually permit one to go into the theory of metric spaces, Cauchy sequences, completeness, etc., necessary for the proof of the general contraction mapping lemma.

The purpose of this note is to give a simple proof of the contraction mapping lemma in $\mathbb{R}^n$ based on a principle familiar to calculus students.

MAX / MIN PRINCIPLE. *A continuous, real-valued function defined on a non-empty, closed, bounded subset of* $\mathbb{R}^n$ *attains its maximum and minimum on that set.*

A proof of this, of course, requires many of the topics one is forced to omit. We feel, however, that our use of this principle to prove the contraction mapping lemma is pedagogically valuable. Early in a calculus course one typically invokes the max/min principle (usually without proof) when discussing max/min problems, and students seem to be reasonably comfortable with it. Its application to the inverse function theorem via the contraction mapping lemma should seem natural and elementary. This approach also serves to introduce students to fixed point techniques, which are powerful tools in many branches of mathematics, particularly differential equations and numerical analysis.

DEFINITION. Let $C \subset \mathbb{R}^n$. A mapping $T: C \to \mathbb{R}^n$ is a *contraction* if there exists a constant α with $0 \leqslant \alpha < 1$ such that

$$\|T(p) - T(q)\| \leqslant \alpha\|p - q\| \quad \text{for all } p, q \in C.$$

THEOREM (Contraction Mapping Lemma on $\mathbb{R}^n$). *Let C be a non-empty, closed subset of $\mathbb{R}^n$, and let $T: C \to C$ be a contraction. Then T has a unique fixed point in C, that is, a point p for which $T(p) = p$.*

Proof. Define $f: C \to \mathbb{R}$ by $f(x) = \|x - T(x)\|$. Note that a zero for f is a fixed point for T. It is easy to show that

$$|f(x) - f(y)| \leqslant (1 + \alpha)\|x - y\|,$$

and so f is continuous. If C is bounded, the max/min principle implies the existence of $p \in C$ such that $f(p)$ is a minimum. Then $f(p) \leqslant f(T(p)) \leqslant \alpha f(p)$. Since $f(p) \geqslant 0$ and $\alpha < 1$, we have $f(p) = 0$.

If C is not bounded, choose $q \in C$ and set

$$\tilde{C} = \{x \in C \mid f(x) \leqslant f(q)\}.$$

If $x \in \tilde{C}$, then

$$\begin{aligned} \|x - q\| &\leqslant \|x - T(x)\| + \|T(x) - T(q)\| + \|T(q) - q\| \\ &\leqslant 2f(q) + \alpha\|x - q\|. \end{aligned}$$

Hence

$$\|x - q\| \leqslant \frac{2f(q)}{1 - \alpha},$$

so $\tilde{C}$ is closed and bounded. From $f(T(p)) \leqslant \alpha f(p)$, it follows that T preserves $\tilde{C}$, and we may proceed as above.

Finally, if $p, q \in C$ are both fixed points of T, then

$$\|p - q\| = \|T(p) - T(q)\| \leqslant \alpha\|p - q\|,$$

and so $\|p - q\| = 0$ and the fixed point is unique.

In closing, we should point out that the contraction mapping lemma is the basis for several iterative numerical methods familiar to undergraduates, e.g., Newton's method. (For a good survey of methods from this point of view, see [2].) In these applications one uses the fact that if x_0 is an arbitrary element of the set on which T is a contraction, then the sequence $x_0, x_1, x_2, \dots$ defined by $x_{n+1} = T(x_n)$ converges geometrically to the fixed point. This property, which is a by-product of the usual proof of the contraction mapping lemma, follows easily from

$$\text{(i)} \qquad f(T(x)) \leqslant \alpha f(x).$$

If p is the fixed point, then

$$\|x - p\| \leqslant \|x - T(x)\| + \|T(x) - T(p)\| \leqslant f(x) + \alpha\|x - p\|,$$

and so

$$\text{(ii)} \qquad \|x - p\| \leqslant \frac{1}{1-\alpha} f(x).$$

Combining (i) and (ii) for the sequence $\{x_n\}$ yields

$$\|x_{n+1} - p\| \leqslant \frac{\alpha^n}{1-\alpha} f(x_0),$$

which shows that the convergence is at least geometric. Inequality (ii) gives a better estimate, however, and shows that the function f gives a convenient test as to when the iteration should stop. If one wishes x_n to be within ε of the fixed point, then the iteration should continue until $f(x_n) < \varepsilon(1 - \alpha)$.

References

1. S. Lang, Differential Manifolds, Addison-Wesley, Reading, MA, 1972.
2. C. H. Wagner, A generic approach to iterative methods, Math. Mag., 55 (1982) 259–273.

MAXIMIZING THE AREA OF A TRAPEZIUM

JOHN P. HOYT
1115 Marietta Avenue, #32, Lancaster, PA 17603

We define a trapezium as a convex quadrilateral with no parallel sides. Generally speaking a trapezium is uniquely determined by five independent conditions. For example, a trapezium is uniquely determined given the lengths of four consecutive sides and the fact that the sum of one pair of opposite angles equals the sum of the other pair. This last condition is the most well-known necessary and sufficient condition that a convex quadrilateral be inscribable in a circle [1, pp. 74 and 88]. Many interesting relationships exist between the parts of an inscribed quadrilateral. We shall have occasion to refer to Brahmagupta's Theorem [2, p. 86]: *If the sides of an inscribed quadrilateral are of lengths a, b, c, and d and if the semi-perimeter is denoted by s, the area of the quadrilateral is given by*

$$A = [(s-a)(s-b)(s-c)(s-d)]^{0.5}.$$

On the other hand a trapezium is not uniquely determined given the lengths of four consecutive sides and the fact that the sum of one pair of opposite sides equals the sum of the other pair. The last condition is the most well-known necessary and sufficient condition that a quadrilateral be circumscribable about a circle [1, p. 89]. There is usually an infinite number of such circles for each given set of sides. However, if in addition to the lengths of the four consecutive sides and the fact that the sums of the opposite pairs of sides are equal, we are given one of the segments on one of the sides made by the point of tangency, then the quadrilateral is uniquely determined and knowing one segment we know them all. If we let the segments made on side a by the point of tangency of a particular inscribed circle be p and q and the segments made on opposite side c by