Machine Precision Derivatives

Keith A. Lewis

December 19, 2019

Abstract

This short note demonstrates a method for computing mixed partial derivatives to machine precision.

Suppose there were a "number" ϵ such that $\epsilon \neq 0$ but $\epsilon^2 = 0$. If f is differentiable at x then, using the Taylor series expansion,

$$f(x + \epsilon) = f(x) + f'(x)\epsilon.$$

For example, if $f(x) = x^2$ then $(x + \epsilon)^2 = x^2 + 2x\epsilon$ so $f'(x) = 2x$.

There is such a "number", the 2×2 matrix $\epsilon = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. No need to compute limits of difference quotients[**?**] or drag in Automatic Differentiation[**?**] machinery or [**?**].

It is possible use this method to compute arbitrary derivatives to machine precision, including mixed partial derivatives.

Functional Calculus

Functions on numbers can be extended to functions on linear operators using a *functional calculus*[**?**]. If $T\colon V \to V$ is a linear operator on the vector space V and p is a polynomial, then $p(T)$ can be defined in the obvious way. If q is a polynomial and $q(T)$ is invertible, then we can define $(p/q)(T) = p(T)q(T)^{-1}$ for appropriate rational functons.

If V is a Banach space we can use power series. If f is sufficiently differentiable, $f(x) = \sum_{n\geq 0} f^{(n)}(0)x^n/n!$ and $f(T) = \sum_{n\geq 0} f^{(n)}(0)T^n/n!$ exists whenever $\|T\|$ is less than the radius of convergence of the series.

The Taylor functional calculus extends this to any function that is holomorphic on a neighborhood of the spectrum of T. $f(T) = \frac{1}{2\pi i}\int_\gamma f(\zeta)(\zeta I - T)^{-1})\,d\zeta$ where γ is a curve surounding the spectrum of T.

Epsilon

Every vector space, V, of dimension n has a linear operator $\epsilon\colon V \to V$ with I, ϵ, …, ϵ^{n-1} independent. E.g., if e_1, …, e_n is a basis of V define $\epsilon e_1 = 0$ and $\epsilon e_j = e_{j-1}$ for $1 < j \leq n$. If $V = \mathbf{R}^n$ with the standard basis, $\epsilon = [\delta_{i,i+1}]$ is the matrix with 1's above the main diagonal and 0's elsewhere. The algebra generated by ϵ is the upper triangular Toepliz matrices.

Univariate Derivatives

Suppose $f\colon \mathbf{R} \to \mathbf{R}$ has derivatives to order n. The Taylor series expansion is

$$f(xI + \epsilon) = \sum_{k=0}^{n-1} \frac{D^k f(x)}{k!} \epsilon^k,$$

since $\epsilon^{n+k} = 0$ for $k \geq 0$. Note $D^k f = f^{(k)}$ is the k-th derivative of f.

For example, if $f(x) = \exp(x) = \sum_{k \geq 0} x^k/k!$ then

$$\begin{aligned}
\exp(xI + \epsilon) &= \sum_{k \geq 0} (xI + \epsilon)^k/k! \\
&= \sum_{k \geq 0} \sum_{j=0}^{k} \binom{k}{j} x^j \epsilon^{k-j}/k! \\
&= \sum_{j \geq 0} \sum_{k \geq j} x^j \epsilon^{k-j}/j!(k-j)! \\
&= \sum_{j \geq 0} \sum_{k \geq 0} x^j \epsilon^k/j!k! \\
&= \sum_{k \geq 0} \sum_{j \geq 0} x^j \epsilon^k/j!k! \\
&= \sum_{k \geq 0} \exp(x) \epsilon^k/k!.
\end{aligned}$$

This shows $D^k \exp(x) = \exp(x)$ for all k.

Multivariate Derivatives

It is also possible to compute mixed derivatives to machine precision. Suppose $f\colon \mathbf{R}^n \to \mathbf{R}$. The Taylor series expansion is

$$f(x+\epsilon) = \sum_{k\geq 0} \sum_{|\alpha|=k} \frac{D^\alpha f(x)}{\alpha!} \epsilon^\alpha,$$

where $x = (x_1, \ldots, x_n)$, $\epsilon = (\epsilon_1, \ldots, \epsilon_n)$, $\alpha = (\alpha_1, \ldots, \alpha_n)$, $|\alpha| = \alpha_1 + \cdots + \alpha_n$, $D^\alpha f(x) = \partial^{|\alpha|} f(x_1, \ldots, x_n)/\partial^{\alpha_1} x_1 \cdots \partial^{\alpha_n} x_n$, $\alpha! = \alpha_1! \cdots \alpha_n!$, and $\epsilon^\alpha = \epsilon_1^{\alpha_1} \cdots \epsilon_n^{\alpha_n}$, a triumph of mathematical notation if there ever was one.

In order to compute mixed derivatives of orders m_j of x_j, $1 \leq j \leq n$, let $\epsilon_j = I \oplus \cdots \oplus \epsilon \oplus \cdots \oplus I$ be the direct of identity operators with ϵ of order m_j in the j-th position. To compute mixed partial derivatives at $x \in \mathbf{R}^n$ evaluate $f(x_1 I + \epsilon_1, \ldots, x_n I + \epsilon_n)$. The coefficient of $\epsilon_1^{\alpha_1} \cdots \epsilon_n^{\alpha_n}$ is the mixed partial corresponding to α.

Computer Implementaton

For the univariate case the algebra generated by ϵ of order n has dimension n. There is an isomorphism with $\mathbf{R}^n$ via $(a_j)_{0\leq j<n} \to \sum_{0\leq j<n} a_j \epsilon^n/j!$. It is convenient to include the factorial in the denominators to map directly to derivatives and for numerical stability.

Obviously, addition and subtraction correspond to the usual vector space operations on $\mathbf{R}^n$. Multiplication is $(a_i)(b_j) = (c_k)$ where $c_k = \sum_{i=0}^{n-1} \binom{k}{i} a_i b_{k-i}$.

Map $f(xI + \epsilon)$ to $(f^{(n)}(x)/n!)$.

Remarks

If f is *analytic* in a neighborhood of $\sigma(T)$, the spectrum of T ...

NOTES

Apache Java math project

http://commons.apache.org/proper/commons-math/

http://commons.apache.org/proper/commons-math/apidocs/org/apache/commons/math4/analysis/differentiati

Computer implemtation only involve $f(xI + \epsilon)$. The rest is taken care of by Toeplitz matrix mulitplication.