Derivative Approximation by Finite Differences

David Eberly
Magic Software, Inc.
http://www.magic-software.com

Created: May 30, 2001
Modified: January 21, 2003 (added bivariate and multivariate discussion)

1 Introduction

This document shows how to approximate derivatives of univariate functions $F(x)$ by finite differences. Given a small value $h > 0$, the d–th order derivative satisfies the following equation where the integer order of error $p > 0$ may be selected as desired,

$$\frac{h^d}{d!}F^{(d)}(x) + O(h^{d+p}) = \sum_{i=i_{\min}}^{i_{\max}} C_i F(x + ih), \tag{1}$$

for some choice of extreme indices $i_{\min}$ and $i_{\max}$ and for some choice of coefficients C_i. The equation becomes an approximation by throwing away the $O(h^{d+p})$ term. The vector $\vec{C} = (C_{i_{\min}}, \ldots, C_{i_{\max}})$ is called the *template* for the approximation. Approximations for the derivatives of multivariate functions are constructed as tensor products of templates for univariate functions.

2 Derivatives of Univariate Functions

Recall from calculus that the following approximations are valid for the derivative of $F(x)$. A *forward difference approximation* is

$$F'(x) = \frac{F(x + h) - F(x)}{h} + O(h), \tag{2}$$

a *backward difference approximation* is

$$F'(x) = \frac{F(x) - F(x - h)}{h} + O(h), \tag{3}$$

and a *centered difference approximation* is

$$F'(x) = \frac{F(x + h) - F(x - h)}{2h} + O(h^2). \tag{4}$$

The approximations are obtained by throwing away the error terms indicated by the O notation. The order of the error for each of these approximations is easily seen from formal expansions as Taylor series about the value x,

$$F(x + h) = F(x) + hF'(x) + \frac{h^2}{2!}F''(x) + \ldots = \sum_{n=0}^{\infty} \frac{h^n}{n!}F^{(n)}(x)$$

and

$$F(x-h) = F(x) - hF'(x) + \frac{h^2}{2!}F''(x) + \ldots = \sum_{n=0}^{\infty}(-1)^n\frac{h^n}{n!}F^{(n)}(x)$$

where $F^{(n)}(x)$ denotes the n–th order derivative of F. The first equation leads to the forward difference $F'(x) = (F(x+h) - F(x))/h + O(h)$. The second equation leads to the backward difference $F'(x) = (F(x) - F(x-h))/h + O(h)$. Both approximations have error $O(h)$. The centered difference is obtained by subtracting the second equation from the first to obtain $(F(x+h) - F(x-h))/(2h) + O(h^2)$.

Higher order approximations to the first derivative can be obtained by using more Taylor series, more terms in the Taylor series, and cleverly weighting the various expansions in a sum. For example,

$$F(x+2h) = \sum_{n=0}^{\infty}\frac{(2h)^n}{n!}F^{(n)}(x) \text{ and } F(x-2h) = \sum_{n=0}^{\infty}(-1)^n\frac{(2h)^n}{n!}F^{(n)}(x)$$

lead to a forward difference approximation with second order error,

$$F'(x) = \frac{-F(x+2h) + 4F(x+h) - 3F(x)}{2h} + O(h^2) \tag{5}$$

to a backward difference approximation with second order error,

$$F'(x) = \frac{3F(x) - 4F(x-h) + F(x-2h)}{2h} + O(h^2), \tag{6}$$

and to a centered difference approximation with fourth order error,

$$F'(x) = \frac{-F(x+2h) + 8F(x+h) - 8F(x-h) + F(x-2h)}{12h} + O(h^4). \tag{7}$$

Higher–order derivatives can be approximated in the same way. For example, a forward difference approximation to $F''(x)$ is

$$F''(x) = \frac{F(x+2h) - 2F(x+h) + F(x)}{h^2} + O(h) \tag{8}$$

and centered difference approximations are

$$F''(x) = \frac{F(x+h) - 2F(x) + F(x-h)}{h^2} + O(h^2) \tag{9}$$

and

$$F''(x) = \frac{-F(x+2h) + 16F(x) - 30F(x) + 16F(x-h) - F(x-2h)}{12h^2} + O(h^4). \tag{10}$$

Each of these formulas is easily verified by expanding the $F(x+ih)$ terms in a formal Taylor series and computing the weighted sums on the right–hand sides. However, of greater interest is to select the order of derivative d and the order of error p and determine the weights C_i for the sum in equation (11). A formal Taylor series for $F(x+ih)$ is

$$F(x+ih) = \sum_{n=0}^{\infty} i^n\frac{h^n}{n!}F^{(n)}(x).$$

Replacing this in equation (11) yields

$$\begin{aligned}
\frac{h^d}{d!}F^{(d)}(x) + O(h^{d+p}) &= \sum_{i=i_{\min}}^{i_{\max}} C_i \sum_{n=0}^{\infty} i^n \frac{h^n}{n!} F^{(n)}(x) \\
&= \sum_{n=0}^{\infty} \left(\sum_{i=i_{\min}}^{i_{\max}} i^n C_i \right) \frac{h^n}{n!} F^{(n)}(x) \\
&= \sum_{n=0}^{d+p-1} \left(\sum_{i=i_{\min}}^{i_{\max}} i^n C_i \right) \frac{h^n}{n!} F^{(n)}(x) + O(h^{d+p}).
\end{aligned}$$

Multiplying by $d!/h^d$, the desired approximation is

$$F^{(d)}(x) = \frac{d!}{h^d} \sum_{n=0}^{d+p-1} \left(\sum_{i=i_{\min}}^{i_{\max}} i^n C_i \right) \frac{h^n}{n!} F^{(n)}(x) + O(h^p). \tag{11}$$

In order for equation (11) to be satisfied, it is necessary that

$$\sum_{i=i_{\min}}^{i_{\max}} i^n C_i = \left\{ \begin{array}{ll} 0, & 0 \le n \le d+p-1 \text{ and } n \ne d \\ 1, & n = d \end{array} \right\}. \tag{12}$$

This is a set of $d+p$ linear equations in $i_{\max} - i_{\min} + 1$ unknowns. If we constrain the number of unknowns to be $d+p$, the linear system has a unique solution. A *forward difference approximation* occurs if we set $i_{\min} = 0$ and $i_{\max} = d+p-1$. A *backward difference approximation* occurs if we set $i_{\max} = 0$ and $i_{\min} = -(d+p-1)$. A *centered difference approximation* occurs if we set $i_{\max} = -i_{\min} = (d+p-1)/2$ where it appears that $d+p$ is necessarily an odd number. As it turns out, p can be chosen to be even regardless of the parity of d and $i_{\max} = \lfloor (d+p-1)/2 \rfloor$.

The table below indicates the choices for d and p, the type of approximation (forward, backward, centered), and the corresponding equation number:

equation	d	p	type	$i_{\min}$	$i_{\max}$
(2)	1	1	forward	0	1
(3)	1	1	backward	-1	0
(4)	1	2	centered	-1	1
(5)	1	2	forward	0	2
(6)	1	2	backward	-2	0
(7)	1	4	centered	-2	2
(8)	2	1	forward	0	2
(9)	2	2	centered	-1	1
(10)	2	4	centered	-2	2

Example 1. Approximate $F^{(3)}(x)$ with a forward difference with error $O(h)$, so $d = 3$ and $p = 1$. We need

$i_{\min} = 0$ and $i_{\max} = 3$. The linear system from equation (12) is

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \\ 0 & 1 & 8 & 27 \end{bmatrix} \begin{bmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

and has solution $(C_0, C_1, C_2, C_3) = (-1, 3, -3, 1)/6$. Equation (11) becomes

$$F^{(3)}(x) = \frac{-F(x) + 3F(x+h) - 3F(x+2h) + F(x+3h)}{h^3} + O(h).$$

Approximate $F^{(3)}(x)$ with a centered difference with error $O(h^2)$, so $d = 3$ and $p = 2$. We need $i_{\max} = -i_{\min} = 2$. The linear system from equation (12) is

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 4 & 1 & 0 & 1 & 4 \\ -8 & -1 & 0 & 1 & 8 \\ 16 & 1 & 0 & 1 & 16 \end{bmatrix} \begin{bmatrix} C_{-2} \\ C_{-1} \\ C_0 \\ C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

and has solution $(C_{-2}, C_{-1}, C_0, C_1, C_2) = (-1, 2, 0, -2, 1)/12$. Equation (11) becomes

$$F^{(3)}(x) = \frac{-F(x-2h) + 2F(x-h) - 2F(x+h) + F(x+2h)}{2h^3} + O(h^2).$$

Finally, approximate with a centered difference with error $O(h^4)$, so $d = 3$ and $p = 4$. We need $i_{\max} = -i_{\min} = 3$. The linear system from equation (12) is

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ 9 & 4 & 1 & 0 & 1 & 4 & 9 \\ -27 & -8 & -1 & 0 & 1 & 8 & 27 \\ 81 & 16 & 1 & 0 & 1 & 16 & 81 \\ -243 & -32 & -1 & 0 & 1 & 32 & 243 \\ 729 & 64 & 1 & 0 & 1 & 64 & 729 \end{bmatrix} \begin{bmatrix} C_{-3} \\ C_{-2} \\ C_{-1} \\ C_0 \\ C_1 \\ C_2 \\ C_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

and has solution $(C_{-3}, C_{-2}, C_{-1}, C_0, C_1, C_2, C_3) = (1, -8, 13, 0, -13, 8, -1)/48$. Equation (11) becomes

$$F^{(3)}(x) = \frac{F(x-3h) - 8F(x-2h) + 13F(x-h) - 13F(x+h) + 8F(x+2h) - F(x+3h)}{8h^3} + O(h^4).$$

EXAMPLE 2. Approximate $F^{(4)}(x)$ with a forward difference with error $O(h)$, so $d = 4$ and $p = 1$. We need $i_{\min} = 0$ and $i_{\max} = 4$. The linear system from equation (12) is

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 1 & 4 & 9 & 16 \\ 0 & 1 & 8 & 27 & 64 \\ 0 & 1 & 16 & 81 & 256 \end{bmatrix} \begin{bmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

and has solution $(C_0, C_1, C_2, C_3, C_4) = (1, -4, 6, -4, 1)/24$. Equation (11) becomes

$$F^{(4)}(x) = \frac{F(x) - 4F(x+h) + 6F(x+2h) - 4F(x+3h) + F(x+4h)}{h^4} + O(h).$$

Approximate $F^{(4)}(x)$ with a centered difference with error $O(h^2)$, so $d = 4$ and $p = 2$. We need $i_{\max} = -i_{\min} = 2$. The linear system from equation (12) is

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 4 & 1 & 0 & 1 & 4 \\ -8 & -1 & 0 & 1 & 8 \\ 16 & 1 & 0 & 1 & 16 \end{bmatrix} \begin{bmatrix} C_{-2} \\ C_{-1} \\ C_0 \\ C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

and has solution $(C_{-2}, C_{-1}, C_0, C_1, C_2) = (1, -4, 6, -4, 1)/24$. Equation (11) becomes

$$F^{(4)}(x) = \frac{F(x-2h) - 4F(x-h) + 6F(x) - 4F(x+h) + F(x+2h)}{h^4} + O(h^2).$$

Finally, approximate with a centered difference with error $O(h^4)$, so $d = 4$ and $p = 4$. We need $i_{\max} = -i_{\min} = 3$. The linear system from equation (12) is

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ 9 & 4 & 1 & 0 & 1 & 4 & 9 \\ -27 & -8 & -1 & 0 & 1 & 8 & 27 \\ 81 & 16 & 1 & 0 & 1 & 16 & 81 \\ -243 & -32 & -1 & 0 & 1 & 32 & 243 \\ 729 & 64 & 1 & 0 & 1 & 64 & 729 \end{bmatrix} \begin{bmatrix} C_{-3} \\ C_{-2} \\ C_{-1} \\ C_0 \\ C_1 \\ C_2 \\ C_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

and has solution $(C_{-3}, C_{-2}, C_{-1}, C_0, C_1, C_2, C_3) = (-1, 12, -39, 56, -39, 12, -1)/144$. Equation (11) becomes

$$F^{(4)}(x) = \frac{-F(x-3h) + 12F(x-2h) - 39F(x-h) + 56F(x) - 39F(x+h) + 12F(x+2h) - F(x+3h)}{6h^4} + O(h^4).$$

3 Derivatives of Bivariate Functions

For functions with more variables, the partial derivatives can be approximated by grouping together all of the same variables and applying the univariate approximation for that group. For example, if $F(x, y)$ is our function, then some partial derivative approximations are

$$\begin{aligned}
f_x(x,y) &\doteq \tfrac{F(x+h,y)-F(x-h,y)}{2h} \\
f_y(x,y) &\doteq \tfrac{F(x,y+k)-F(x,y-k)}{2k} \\
f_{xx}(x,y) &\doteq \tfrac{F(x+h,y)-2f(x,y)+F(x-h,y)}{h^2} \\
f_{yy}(x,y) &\doteq \tfrac{F(x,y+k)-2f(x,y)+F(x,y-k)}{k^2} \\
f_{xy}(x,y) &\doteq \tfrac{F(x+h,y+k)-F(x+h,y-k)-F(x-h,y+k)+F(x-h,y-k)}{4hk}
\end{aligned}$$

Each of these can be verified in the limit, the x–derivatives by taking the limit as h approaches zero, the y–derivatives by taking the limit as y approaches zero, and the mixed second–order derivative by taking the limit as both h and k approach zero.

The derivatives F_x, F_y, F_{xx}, and F_{yy} just use the univariate approximation formulas. The mixed derivative requires slightly more work. The important observation is that the approximation for F_{xy} is obtained by applying the x–derivative approximation for F_x, then applying the y–derivative approximation to the previous approximation. That is,

$$\begin{aligned}
f_{xy}(x,y) &\doteq \tfrac{F(x+h,y)-F(x-h,y)}{2h} \\
&\doteq \tfrac{\frac{F(x+h,y+k)-F(x-h,y+k)}{2h} - \frac{F(x+h,y-k)-F(x-h,y-k)}{2h}}{2k} \\
&= \tfrac{F(x+h,y+k)-F(x+h,y-k)-F(x-h,y+k)+F(x-h,y-k)}{4hk}
\end{aligned}$$

The approximation implied by equation (1) may be written as

$$\frac{h^m}{m!}\frac{d^m}{dx^m}F(x) \doteq \sum_{i=i_{\min}}^{i_{\max}} C_i^{(m)} F(x+ih), \tag{13}$$

The inclusion of the superscript on the C coefficients is to emphasize that those coefficients are constructed for each order m. For bivariate functions, we can use the natural extension of equation (13) by applying the approximation in x first, then applying the approximation in y to that approximation, just as in our example of F_{xy}.

$$\begin{aligned}
\tfrac{k^n}{n!}\tfrac{\partial^n}{\partial y^n}\tfrac{h^m}{m!}\tfrac{\partial^m}{\partial x^m}F(x,y) &\doteq \tfrac{k^n}{n!}\tfrac{\partial^n}{\partial y^n}\textstyle\sum_{i=i_{\min}}^{i_{\max}} C_i^{(m)} F(x+ih,y) \\
&\doteq \textstyle\sum_{i=i_{\min}}^{i_{\max}}\sum_{j=j_{\min}}^{j_{\max}} C_i^{(m)} C_j^{(n)} F(x+ih,y+jk) \\
&= \textstyle\sum_{i=i_{\min}}^{i_{\max}}\sum_{j=j_{\min}}^{j_{\max}} C_{i,j}^{(m,n)} F(x+ih,y+jk)
\end{aligned} \tag{14}$$

where the last equality defines

$$C_{i,j}^{(m,n)} = C_i^{(m)} C_j^{(n)}$$

The coefficients for the bivariate approximation are just the tensor product of the coefficients for each of the univariate approximations.

4 Derivatives of Multivariate Functions

The approximation concept extends to any number of variables. Let $(x_1, \ldots, x_n)$ be those variables and let $F(x_1, \ldots, x_n)$ be the function to approximate. The approximation is

$$\left(\frac{h_1^{m_1}}{m_1!} \frac{\partial^{m_1}}{\partial x_1^{m_1}} \cdots \frac{h_n^{m_n}}{m_n!} \frac{\partial^{m_n}}{\partial x_1^{m_n}} \right) F(x_1, \ldots, x_n) \doteq \sum_{i_1=i_1^{\min}}^{i_1^{\max}} \cdots \sum_{i_n=i_n^{\min}}^{i_n^{\max}} C_{(i_1,\ldots,i_n)}^{(m_1,\ldots,m_n)} F(x_1 + i_1 h_1, \ldots, x_n + i_n h_n) \quad (15)$$

where

$$C_{(i_1,\ldots,i_n)}^{(m_1,\ldots,m_n)} = C_{i_1}^{(m_1)} \cdots C_{i_n}^{(m_n)}$$

a tensor product of the coefficients of the n univariate approximations.