Derivative Approximation by Finite Differences

David Eberly, Geometric Tools, Redmond WA 98052
https://www.geometrictools.com/

Created: May 30, 2001
Last Modified: February 1, 2020

Contents

1 Introduction

This document shows how to approximate derivatives of functions $F : \mathbb{R}^n \to \mathbb{R}$ using finite differences. The independent variables are $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and the dependent variable is $y = F(\mathbf{x})$. The function is said to be *univariate* when $n = 1$, *bivariate* when $n = 2$, or generally *multivariate* for $n > 1$.

The derivative of order $m \geq 0$ for univariate $y = F(x)$ is represented by $F^{(m)}(x)$. The function itself occurs when $m = 0$.

A bivariate function $y = F(x_1, x_2)$ can be differentiated $m_1 \geq 0$ times with respect to x_1 and $m_2 \geq 0$ time with respect to x_2. The order of the derivative is $m_1 + m_2$ and the derivative is represented by $F^{(m_1,m_2)}(x_1, x_2)$. The function itself occurs when $(m_1, m_2) = (0, 0)$.

Generally, a multivariate function $y = F(x_1, \ldots, x_n)$ can be differentiated $m_i \geq 0$ times with respect to x_i for each i. The order of the derivative is $\sum_{i=1}^{n} m_i$ and the derivative is represented by $F^{(m_1,\ldots,m_n)}(x_1, \ldots, x_n)$. The function itself occurs when $m_i = 0$ for all i. Multiindex notation can be used to obtain more concise notation. Define $\mathbf{m} = (m_1, \ldots, m_n)$ be a multiindex with nonnegative components. The zero multiiindex is $\mathbf{0}$ which has all zero components. Define $\|\mathbf{m}\| = \sum_{i=1}^{n} m_i$. The derivative is represented by $F^{(\mathbf{m})}(\mathbf{x})$ and has order $\|\mathbf{m}\|$. The function itself occurs when $\mathbf{m} = \mathbf{0}$.

2 Derivative Approximations for Univariate Functions

Given a small number $h > 0$, the derivative of order m for a univariate function satisfies the following equation,

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

where $p > 0$ and where the extreme indices $i_{\min}$ and $i_{\max}$ are chosen subject to the constraints $i_{\min} < i_{\max}$ and $i_{\max} - i_{\min} + 1 = m + p$. A formal Taylor series for $F(x + ih)$ is

$$F(x + ih) = \sum_{k=0}^{\infty} i^k \frac{h^k}{k!} F^{(k)}(x) = \sum_{k=0}^{m+p-1} i^k \frac{h^k}{k!} F^{(k)}(x) + O(h^{m+p}) \tag{2}$$

The second equality is valid for any $m > 0$ and $p > 0$. Replacing this in equation (1),

$$\begin{array}{rcl} \frac{h^m}{m!} F^{(m)}(x) & = & \sum_{i=i_{\min}}^{i_{\max}} C_i \sum_{k=0}^{\infty} i^k \frac{h^k}{k!} F^{(k)}(x) + O(h^{m+p}) \\ & = & \sum_{k=0}^{\infty} \left(\sum_{i=i_{\min}}^{i_{\max}} i^k C_i \right) \frac{h^k}{k!} F^{(k)}(x) + O(h^{m+p}) \\ & = & \sum_{k=0}^{m+p-1} \left(\sum_{i=i_{\min}}^{i_{\max}} i^k C_i \right) \frac{h^k}{k!} F^{(k)}(x) + O(h^{m+p}) \end{array} \tag{3}$$

In equation (3), the only term in the sum on the right-hand side that contains $(h^m/m!)F^{(m)}(x)$ occurs when $k = m$, so the coefficient of that term must be 1. The other terms must vanish, so the coefficients of those terms must be 0. Therefore, it is necessary that

$$\sum_{i=i_{\min}}^{i_{\max}} i^k C_i = \begin{cases} 0, & 0 \leq k \leq m + p - 1 \text{ and } k \neq m \\ 1, & k = m \end{cases} \tag{4}$$

where 0^0 is defined to be 1 (when $i = 0$ and $k = 0$). This is a set of $m + p$ linear equations in $i_{\max} - i_{\min} + 1$ unknowns. If the number of unknowns is $m + p$, obtained by constraining $i_{\max} - i_{\min} + 1 = m + p$, the linear system has a unique solution. Define $\mathbf{C} = (C_{i_{\min}}, \ldots C_{i_{\max}})$, which is formatted as an $(m + p) \times 1$ vector in linear algebraic operations. Define W to be the $(m + p) \times (m + p)$ matrix of coefficients of the linear system. Define $\mathbf{e}$ to be the zero-indexed $(m + p) \times 1$ vector whose mth component is 1 and all other components are 0. The linear system is $W\mathbf{C} = \mathbf{e}$ and can be solved for $\mathbf{C} = W^{-1}\mathbf{e}$.

The derivative approximation is obtained by solving for $F^{(m)}(x)$ in equation (1),

$$F^{(m)}(x) = \frac{m!}{h^m} \sum_{i=i_{\min}}^{i_{\max}} C_i F(x + ih) + O(h^p) = \sum_{i=i_{\min}}^{i_{\max}} \frac{R_i F(x + ih)}{h^m} + O(h^p) \tag{5}$$

where the R_i are rational numbers. Each rational number can be written in canonical form, $R_i = \hat{N}_i / \hat{D}_i$, where the greatest common divisior of $\hat{N}_i$ and $\hat{D}_i$ is 1. Compute the least common multiple D of the denominators. The rational numbers are then $R_i = N_i / D$, where D and $N_i = \hat{N}_i D / \hat{D}_i$ are integers. The final form of the derivative approximation is

$$F^{(m)}(x) \doteq \sum_{i=i_{\min}}^{i_{\max}} \frac{N_i F(x + ih)}{D h^m} + O(h^p) \tag{6}$$

A *forward-difference approximation* occurs when $i_{\min} \geq 0$. A *backward-difference approximation* occurs when $i_{\max} \leq 0$. A *mixed-difference approximation* occurs when $i_{\min} < 0 < i_{\max}$.

A special case of a mixed-difference approximation is a *centered-difference approximation*, where $i_{\max} = -i_{\min}$. The number of unknowns $m + p$ must be odd so that $\ell = (m + p - 1)/2$ is an integer. The extreme indices are $i_{\min} = -\ell$ and $i_{\max} = \ell$. The linear system $W\mathbf{C} = \mathbf{e}$ can be reduced to linear subsystems whose solutions are combined to determine $\mathbf{C}$.

When m is odd, $C_i = -C_{-i}$ for all i, which forces $C_0 = 0$. When m is even, $C_i = C_{-i}$ for all i. To see this, define $s_i = C_i + C_{-i}$ and $d_i = C_i - C_{-i}$ for $1 \leq i \leq \ell$. Define $\mathbf{s}$ to be the $\ell \times 1$ vector whose components are the s_i and define $\mathbf{d}$ to be the $\ell \times 1$ vector whose components are the d_i. The linear system has an equation $\sum_{i=-\ell}^{\ell} C_i = 0$. It also has 2ℓ equations for powers $k > 0$, where each equation has coefficient 0 for the variable C_0. Half of the equations involve even powers $k = 2r$ and half of the equations involve odd powers $k = 2r - 1$ for $1 \leq r \leq \ell$. The left-hand side terms in the equations are combined to form

$$\sum_{i=-\ell}^{\ell} i^{2r} C_i = \sum_{i=1}^{\ell} i^{2r} S_i, \quad \sum_{i=-\ell}^{\ell} i^{2r-1} C_i = \sum_{i=1}^{\ell} i^{2r} D_i \tag{7}$$

The linear system $W\mathbf{C} = \mathbf{e}$ reduces to the equation $\sum_{i=-\ell}^{\ell} C_i = 0$ and two linear subsystems, each with ℓ equations in ℓ unknowns. Define U to be the $\ell \times \ell$ matrix of coefficients for the even-power equations, define V be the $\ell \times \ell$ matrix of coefficients for the odd-power equations and define $\mathbf{f}$ to have 1 in the component corresponding to m and all other components 0. If m is odd, the subsystems are $U\mathbf{s} = \mathbf{0}$ and $V\mathbf{d} = \mathbf{f}$. The first subsystem has solution $\mathbf{s} = \mathbf{0}$, which implies $C_i + C_{-i} = 0$ for all i. In particular, $C_0 = 0$. The second subsystem has solution $\mathbf{d} = V^{-1}\mathbf{f}$. The outcome is that all the C_i are determined. If m is even, the subsystems are $U\mathbf{s} = \mathbf{f}$ and $V\mathbf{d} = \mathbf{0}$. The second subsystem has solution $\mathbf{d} = \mathbf{0}$, which implies $C_i - C_{-i} = 0$ for all $i > 0$. The first subsystem has solution $\mathbf{s} = U^{-1}\mathbf{f}$. The outcome is that all the C_i are determined for $i > 0$. The equation $\sum_{i=-\ell}^{\ell} C_i = 0$ determines $C_0 = -2\sum_{i=1}^{\ell} C_i$.

Another observation is that when m is even and p is odd, the order of the centered-difference approximation is actually $p+1$. To see this, it was proved in the previous paragraph that $C_i = C_{-i}$ for all i. The last row of W is generated by the power $m+p-1$. The implication of these two facts is that $\sum_{i=-\ell}^{\ell} i^{m+p} C_i = 0$. The $O(h^p)$ terms in the Taylor series used to construct the approximation also cancel, so the actual error is $O(h^{p+1})$.

3 Derivative Approximations for Bivariate Functions

Given small numbers $h_1 > 0$ and $h_2 > 0$ and derivative orders $m_1 \geq 0$ and $m_2 \geq 0$ for a bivariate function, a derivative approximation is provided by the following equation,

$$F^{(m_1,m_2)}(x_1,x_2) \doteq \left(\frac{m_1!}{h_1^{m_1}} \frac{m_2!}{h_2^{m_2}}\right) \sum_{i_1=i_{1,\min}}^{i_{1,\max}} \sum_{i_2=i_{2,\min}}^{i_{2,\max}} (C_{1,i_1} C_{2,i_2})\, F(x_1 + i_1 h_1, x_2 + i_2 h_2) + \max_{1\leq j \leq 2} \left\{ O\left(h_j^{p_j}\right) \right\} \tag{8}$$

The coefficients are an outer product of $\mathbf{C}_1 = (C_{1,i_{1,\min}}, \ldots, C_{1,i_{1,\max}})$ and $\mathbf{C}_2 = (C_{2,i_{2,\min}}, \ldots, C_{2,i_{2,\max}})$. Each vector $\mathbf{C}_j$ consists of coefficients computed using the algorithm for estimation of derivatives of univariate functions; the j index corresponds to x_j.

4 Derivative Approximations for Multivariate Functions

Given small numbers $h_j > 0$ and derivative orders $m_j \geq 0$ for $1 \leq j \leq n$ for a multivariate function, a derivative approximation is provided by the following equation written using multiindex notation,

$$F^{(\mathbf{m})}(\mathbf{x}) \doteq \frac{\mathbf{m}!}{\mathbf{h}^{m}} \sum_{\mathbf{i}=\mathbf{i}_{\min}}^{\mathbf{i}_{\max}} \mathbf{K}_{\mathbf{i}} F(\mathbf{x} + \mathbf{ih}) + \max_{1\leq j\leq n} \left\{ O\left(h_j^{p_j}\right) \right\} \tag{9}$$

The multiindices are $\mathbf{m} = (m_1, \ldots, m_n)$, $\mathbf{i}_{\min} = (i_{1,\min}, \ldots, i_{n,\min})$, $\mathbf{i}_{\max} = (i_{1,\max}, \ldots, i_{n,\max})$ and $\mathbf{i} = (i_1, \ldots, i_n)$. Real-valued vectors are $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{h} = (h_1, \ldots, h_n)$. The expression $\mathbf{m}!$ is the product of the factorials of the components of $\mathbf{m}$, namely, $m_1! \cdots m_n!$. The expression $\mathbf{h}^{\mathbf{m}}$ is the product $h_1^{m_1} \cdots h_n^{m_n}$. The product $\mathbf{ih}$ is performed componentwise; that is, $\mathbf{ih} = (i_1 h_1, \ldots, i_n h_n)$ which is then considered to be a real-valued vector that can be added to $\mathbf{x}$. The tensor $\mathbf{K} = \mathbf{C}_1 \otimes \cdots \otimes \mathbf{C}_n = \bigotimes_{j=1}^{n} \mathbf{C}_j$ is an outer product of the vectors $\mathbf{C}_j$, each vector corresponding to the first-order derivative approximation for x_j. The multiindexed component is $\mathbf{K}_{\mathbf{i}} = C_{1,i_1} \cdots C_{n,i_n}$. Observe the similarity between equation (5) and equation (9).

5 Table of Approximations for First-Order Derivatives

Table 1 contains the approximations constructed for first-order derivatives ($m = 1$). The format of the approximations is

$$F^{(1)}(x) = \sum_{i=i_{\min}}^{i_{\max}} \frac{N_i F(x+ih)}{Dh} + O(h^p) \tag{10}$$

where N_i and D are integers. The s-column is the size of the linear system that was solved to compute the C_i. The p-column refers to the error term $O(h^p)$. The t-column refers to the type of the approximation: f for a forward difference ($i_{\min} \geq 0$), b for a backward difference ($i_{\max} \leq 0$), c for a centered difference ($i_{\max} = -i_{\min}$) or m for a mixed difference ($i_{\min} < 0 < i_{\max}$, $i_{\max} \neq -i_{\min}$). The D-column and the N_i-columns are the denominator and the numerators for the rational coefficients.

Table 1. The approximations for first-order derivatives.

s	p	$i_{\min}$	$i_{\max}$	t	D	N_{-4}	N_{-3}	N_{-2}	N_{-1}	N_0	N_1	N_2	N_3	N_4
2	1	0	1	f	1					−1	1			
2	1	−1	0	b	1				−1	1				
3	2	0	2	f	2					−3	4	−1		
3	2	−1	1	c	2				−1	0	1			
3	2	−2	0	b	2			1	−4	3				
4	3	0	3	f	6					−11	18	−9	2	
4	3	−1	2	m	6				−2	−3	6	−1		
4	3	−2	1	m	6			1	−6	3	2			
4	3	−3	0	b	6		−2	9	−18	11				
5	4	0	4	f	12					−25	48	−36	16	−3
5	4	−1	3	m	12				−3	−10	18	−6	1	
5	4	−2	2	c	12			1	−8	0	8	−1		
5	4	−3	1	m	12		−1	6	−18	10	3			
5	4	−4	0	b	12	3	−16	36	−48	25				

6 Table of Approximations for Second-Order Derivatives

Table 2 contains the approximations constructed for second-order derivatives ($m = 2$). The format of the approximations is

$$F^{(2)}(x) = \sum_{i=i_{\min}}^{i_{\max}} \frac{N_i F(x+ih)}{Dh^2} + O(h^p) \tag{11}$$

where N_i and D are integers. The s-column is the size of the linear system that was solved to compute the C_i. The p-column refers to the error term $O(h^p)$. The t-column refers to the type of the approximation: f for a forward difference ($i_{\min} \geq 0$), b for a backward difference ($i_{\max} \leq 0$), c for a centered difference ($i_{\max} = -i_{\min}$) or m for a mixed difference ($i_{\min} < 0 < i_{\max}$, $i_{\max} \neq -i_{\min}$). The D-column and the N_i-columns are the denominator and the numerators for the rational coefficients.

Table 2. The approximations for second-order derivatives.

s	p	$i_{\min}$	$i_{\max}$	t	D	N_{-5}	N_{-4}	N_{-3}	N_{-2}	N_{-1}	N_0	N_1	N_2	N_3	N_4	N_5
3	1	0	2	f	1						1	−2	1			
3	2	−1	1	c	1					1	−2	1				
3	1	−2	0	b	1				1	−2	1					
4	2	0	3	f	1						2	−5	4	−1		
4	2	−1	2	m	1					1	−2	1	0			
4	2	−2	1	m	1				0	1	−2	1				
4	2	−3	0	b	1			−1	4	−5	2					
5	3	0	4	f	12						35	−104	114	−56	11	
5	3	−1	3	m	12					11	−20	6	4	−1		
5	4	−2	2	c	12				−1	16	−30	16	−1			
5	3	−3	1	m	12			−1	4	6	−20	11				
5	3	−4	0	b	12		11	−56	114	−104	35					
6	4	0	5	f	12						45	−154	214	−156	61	−10
6	4	−1	4	m	12					10	−15	−4	14	−6	1	
6	4	−2	3	m	12				−1	16	−30	16	1	0		
6	4	−3	2	m	12			0	−1	16	−30	16	1			
6	4	−4	1	m	12		1	−6	14	−4	−15	10				
6	4	−5	0	b	12	−10	61	−156	214	−154	45					

7 Table of Approximations for Third-Order Derivatives

Table 3 contains the approximations constructed for third-order derivatives ($m = 3$). The format of the approximations is

$$F^{(3)}(x) = \sum_{i=i_{\min}}^{i_{\max}} \frac{N_i F(x + ih)}{Dh^3} + O(h^p) \tag{12}$$

where N_i and D are integers. The s-column is the size of the linear system that was solved to compute the C_i. The p-column refers to the error term $O(h^p)$. The t-column refers to the type of the approximation: f for a forward difference ($i_{\min} \geq 0$), b for a backward difference ($i_{\max} \leq 0$), c for a centered difference ($i_{\max} = -i_{\min}$) or m for a mixed difference ($i_{\min} < 0 < i_{\max}$, $i_{\max} \neq -i_{\min}$). The D-column and the N_i-columns are the denominator and the numerators for the rational coefficients.

Table 3. The approximations for third-order derivatives.

s	p	$i_{\min}$	$i_{\max}$	t	D	N_{-6}	N_{-5}	N_{-4}	N_{-3}	N_{-2}	N_{-1}	N_0	N_1	N_2	N_3	N_4	N_5	N_6
4	1	0	3	f	1							−1	3	−3	1			
4	1	−1	2	m	1						−1	3	−3	1				
4	1	−2	1	m	1					−1	3	−3	1					
4	1	−3	0	b	1				−1	3	−3	1						
5	2	0	4	f	2							−5	18	−24	14	−3		
5	2	−1	3	m	2						−3	10	−12	6	−1			
5	2	−2	2	c	2					−1	2	0	−2	1				
5	2	−3	1	m	2				1	−6	12	−10	3					
5	2	−4	0	b	2			3	−14	24	−18	5						
6	3	0	5	f	4							−17	71	−118	98	−41	7	
6	3	−1	4	m	4						−7	25	−34	22	−7	1		
6	3	−2	3	m	4					−1	−1	10	−14	7	−1			
6	3	−3	2	m	4				1	−7	14	−10	1	1				
6	3	−4	1	m	4			−1	7	−22	34	−25	7					
6	3	−5	0	b	4		−7	41	−98	118	−71	17						
7	4	0	6	f	8							−49	232	−461	496	−307	104	−15
7	4	−1	5	m	8						−15	56	−83	64	−29	8	−1	
7	4	−2	4	m	8					−1	−8	35	−48	29	−8	1		
7	4	−3	3	c	8				1	−8	13	0	−13	8	−1			
7	4	−4	2	m	8			−1	8	−29	48	−35	8	1				
7	4	−5	1	m	8		1	−8	29	−64	83	−56	15					
7	4	−6	0	b	8	15	−104	307	−496	461	−232	49						

8 Table of Approximations for Fourth-Order Derivatives

Table 4 contains the approximations constructed for fourth-order derivatives ($m = 4$). The format of the approximations is

$$F^{(4)}(x) = \sum_{i=i_{\min}}^{i_{\max}} \frac{N_i F(x+ih)}{Dh^4} + O(h^p) \tag{13}$$

where N_i and D are integers. The s-column is the size of the linear system that was solved to compute the C_i. The p-column refers to the error term $O(h^p)$. The t-column refers to the type of the approximation: f for a forward difference ($i_{\min} \geq 0$), b for a backward difference ($i_{\max} \leq 0$), c for a centered difference ($i_{\max} = -i_{\min}$) or m for a mixed difference ($i_{\min} < 0 < i_{\max}$, $i_{\max} \neq -i_{\min}$). The D-column and the N_i-columns are the denominator and the numerators for the rational coefficients.

Table 4. The approximations for fourth-order derivatives.

s	p	$i_{\min}$	$i_{\max}$	t	D	N_{-7}	N_{-6}	N_{-5}	N_{-4}	N_{-3}	N_{-2}	N_{-1}	N_0	N_1	N_2	N_3	N_4	N_5	N_6	N_7
5	1	0	4	f	1								1	−4	6	−4	1			
5	1	−1	3	m	1							1	−4	6	−4	1				
5	2	−2	2	c	1						1	−4	6	−4	1					
5	1	−3	1	m	1					1	−4	6	−4	1						
5	1	−4	0	b	1				1	−4	6	−4	1							
6	2	0	5	f	1								3	−14	26	−24	11	−2		
6	2	−1	4	m	1							2	−9	16	−14	6	−1			
6	2	−2	3	m	1						1	−4	6	−4	1	0				
6	2	−3	2	m	1					0	1	−4	6	−4	1					
6	2	−4	1	m	1				−1	6	−14	16	−9	2						
6	2	−5	0	b	1			−2	11	−24	26	−14	3							
7	3	0	6	f	6								35	−186	411	−484	321	−114	17	
7	3	−1	5	m	6							17	−84	171	−184	111	−36	5		
7	3	−2	4	m	6						5	−18	21	−4	−9	6	−1			
7	4	−3	3	c	6					−1	12	−39	56	−39	12	−1				
7	3	−4	2	m	6				−1	6	−9	−4	21	−18	5					
7	3	−5	1	m	6			5	−36	111	−184	171	−84	17						
7	3	−6	0	b	6		17	−114	321	−484	411	−186	35							
8	4	0	7	f	6								56	−333	852	−1219	1056	−555	164	−21
8	4	−1	6	m	6							21	−112	255	−324	251	−120	33	−4	
8	4	−2	5	m	6						4	−11	0	31	−44	27	−8	1		
8	4	−3	4	m	6					−1	12	−39	56	−39	12	−1	0			
8	4	−4	3	m	6				0	−1	12	−39	56	−39	12	−1				
8	4	−5	2	m	6			1	−8	27	−44	31	0	−11	4					
8	4	−6	1	m	6		−4	33	−120	251	−324	255	−112	21						
8	4	−7	0	b	6	−21	164	−555	1056	−1219	852	−333	56							