# Taylor Expansion

n.douglas

February 2023

## 1 Taylor Expansion

Any real infinitely differentiable function $f(x)$ at a given point $a$ can be expressed in terms of its Taylor series:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \frac{1}{6}f'''(a)(x-a)^3 + ....$$

or equivalently the function $f(x+h)$ at the point $x$:

$$f(x+h) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!}h^n = f(x) + f'(x)h + \frac{1}{2}f''(x)h^2 + \frac{1}{6}f'''(x)h^3 + ....$$

This allows us to approximate $f(x)$ with a polynomial up to a chosen order. For example, $f(x)$ can be approximated at $x = a$ as the polynomial of order 2, i.e. the quadratic:

$$f(x+h) \approx f(x) + f'(x)h + \frac{1}{2}f''(x)h^2$$

In many applications, including data assimilation, we often require the linear approximation:

$$f(x+h) \approx f(x) + f'(x)h$$

### 1.1 Multivariable Taylor Expansion

To extend the above concept of a linear approximation to functions of two variables in order to approximate $f(x+a, y+b)$ at the point $(x, y)$, we require partial derivatives:

$$f(x+\delta x, y+\delta y) \approx f(x,y) + f_x(x,y)\delta x + f_y(x,y)\delta y$$

where $f_x(x,y)$ and $f_y(x,y)$ are the partial derivatives of $f(x,y)$ with respect to $x$ and $y$ respectively. It follows that a vector of functions:

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \end{pmatrix}$$

can be approximated at a point $\mathbf{x} = (x, y)$ in a similar way:

$$\mathbf{f}(\mathbf{x}+\delta\mathbf{x}) = \begin{pmatrix} f_1(x+\delta x, y+\delta y) \\ f_2(x+\delta x, y+\delta y) \end{pmatrix} \approx \begin{pmatrix} f_1(x,y) + \frac{\partial f_1}{\partial x}(x,y)\delta x + \frac{\partial f_1}{\partial y}(x,y)\delta y \\ f_2(x,y) + \frac{\partial f_2}{\partial x}(x,y)\delta x + \frac{\partial f_2}{\partial y}(x,y)\delta y \end{pmatrix} = \mathbf{f}(\mathbf{x}) + \mathbb{J}(\mathbf{f}, \mathbf{x})\delta\mathbf{x}$$

where $\delta\mathbf{x} = \begin{pmatrix} \delta x \\ \delta y \end{pmatrix}$ and $\mathbb{J}(\mathbf{f},\mathbf{x})$ is the Jacobian:

$$\mathbb{J}(\mathbf{f},\mathbf{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix}$$

The linear approximation

$$\mathbf{f}(\mathbf{x}+\delta\mathbf{x}) \approx \mathbf{f}(\mathbf{x}) + \mathbb{J}(\mathbf{f},\mathbf{x})\delta\mathbf{x}$$

also applies for higher dimensions where

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \delta\mathbf{x} = \begin{pmatrix} \delta x_1 \\ \delta x_2 \\ \vdots \\ \delta x_n \end{pmatrix}, \quad \mathbb{J}(\mathbf{f},\mathbf{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Higher order Taylor expansions, for the quadratic approximation for instance, would require calculation of the Hessian in addition to the Jacobian.

## 1.2 Taylor Expansion for Compound Functions

Consider the compound function $\mathbf{f} \circ \mathbf{g}(\mathbf{x}) = \mathbf{f}(\mathbf{g}(\mathbf{x}))$ for $\mathbf{f} : \mathbb{R}^n \mapsto \mathbb{R}^m$ and $\mathbf{g} : \mathbb{R}^n \mapsto \mathbb{R}^n$. The Taylor expansion is given by

$$\mathbf{f}(\mathbf{g}(\mathbf{x}+\delta\mathbf{x})) = \mathbf{f}(\mathbf{g}(\mathbf{x})) + \mathbb{J}(\mathbf{f}\circ\mathbf{g},\mathbf{x})\delta\mathbf{x} = \mathbf{f}(\mathbf{g}(\mathbf{x})) + \mathbb{J}(\mathbf{f},\mathbf{g})\mathbb{J}(\mathbf{g},\mathbf{x})\delta\mathbf{x}$$

where we have used

$$\begin{aligned}
\mathbb{J}(\mathbf{f}\circ\mathbf{g},\mathbf{x}) &= \begin{pmatrix} \frac{\partial(\mathbf{f}\circ\mathbf{g})_1}{\partial x_1} & \frac{\partial(\mathbf{f}\circ\mathbf{g})_1}{\partial x_2} & \cdots & \frac{\partial(\mathbf{f}\circ\mathbf{g})_1}{\partial x_n} \\ \frac{\partial(\mathbf{f}\circ\mathbf{g})_2}{\partial x_1} & \frac{\partial(\mathbf{f}\circ\mathbf{g})_2}{\partial x_2} & \cdots & \frac{\partial(\mathbf{f}\circ\mathbf{g})_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial(\mathbf{f}\circ\mathbf{g})_m}{\partial x_1} & \frac{\partial(\mathbf{f}\circ\mathbf{g})_m}{\partial x_2} & \cdots & \frac{\partial(\mathbf{f}\circ\mathbf{g})_m}{\partial x_n} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial f_1}{\partial g_1}\frac{\partial g_1}{\partial x_1} + ... + \frac{\partial f_1}{\partial g_n}\frac{\partial g_n}{\partial x_1} & \frac{\partial f_1}{\partial g_1}\frac{\partial g_1}{\partial x_2} + ... + \frac{\partial f_1}{\partial g_n}\frac{\partial g_n}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial g_1}\frac{\partial g_1}{\partial x_2} + ... + \frac{\partial f_1}{\partial g_n}\frac{\partial g_n}{\partial x_n} \\ \frac{\partial f_2}{\partial g_1}\frac{\partial g_1}{\partial x_1} + ... + \frac{\partial f_2}{\partial g_n}\frac{\partial g_n}{\partial x_1} & \frac{\partial f_2}{\partial g_1}\frac{\partial g_1}{\partial x_2} + ... + \frac{\partial f_2}{\partial g_n}\frac{\partial g_n}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial g_1}\frac{\partial g_1}{\partial x_2} + ... + \frac{\partial f_2}{\partial g_n}\frac{\partial g_n}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial g_1}\frac{\partial g_1}{\partial x_1} + ... + \frac{\partial f_m}{\partial g_n}\frac{\partial g_n}{\partial x_1} & \frac{\partial f_m}{\partial g_1}\frac{\partial g_1}{\partial x_2} + ... + \frac{\partial f_m}{\partial g_n}\frac{\partial g_n}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial g_1}\frac{\partial g_1}{\partial x_2} + ... + \frac{\partial f_m}{\partial g_n}\frac{\partial g_n}{\partial x_n} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix} \\
&= \mathbb{J}(\mathbf{f},\mathbf{g})\mathbb{J}(\mathbf{g},\mathbf{x}) \qquad (1)
\end{aligned}$$

It follows that for any number of compounded functions

$$\mathbf{f}\circ\mathbf{g}_1\circ ... \circ\mathbf{g}_p(\mathbf{x}+\delta\mathbf{x}) = \mathbf{f}\circ\mathbf{g}_1\circ ... \circ\mathbf{g}_p(\mathbf{x}) + \mathbb{J}(\mathbf{f},\mathbf{g}_1)\mathbb{J}(\mathbf{g}_1,\mathbf{g}_2)...\mathbb{J}(\mathbf{g}_p,\mathbf{x})\delta\mathbf{x}$$

where $\mathbf{f} : \mathbb{R}^n \mapsto \mathbb{R}^m$ and $\mathbf{g}_i : \mathbb{R}^n \mapsto \mathbb{R}^n$ for $i = 1, .., p$.