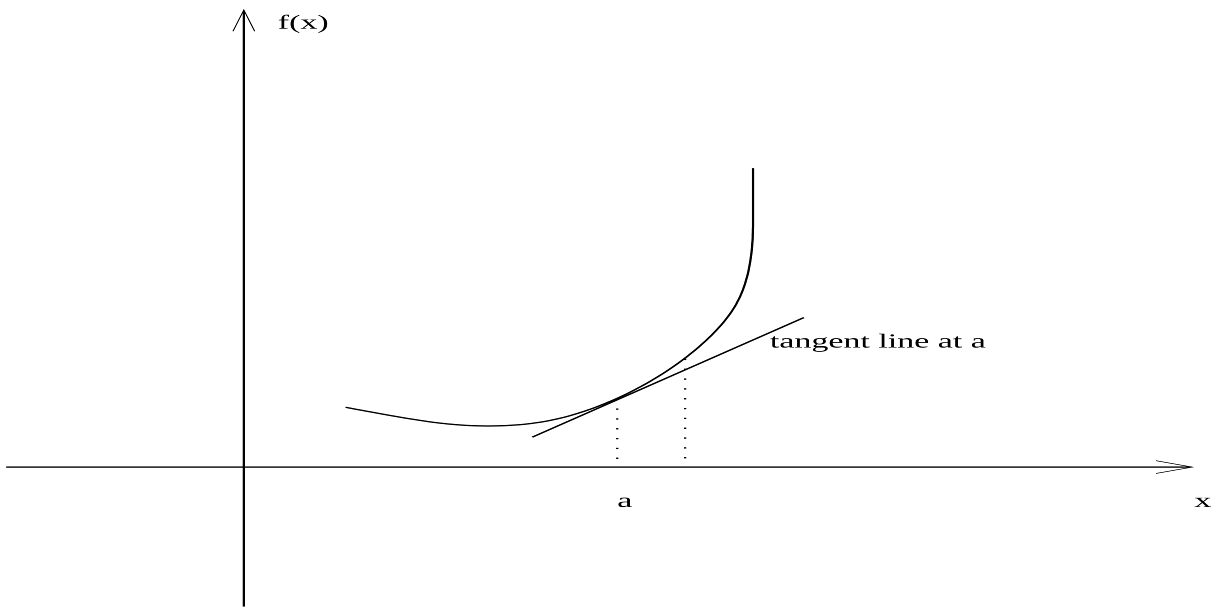
Taylor Series

Linear approximation: Linear approximation is to approximate a general function using a linear function. Given a differentiable scalar function $f(x)$ of the real variable x, the linear approximation of the function at point a, as shown in the Figure below, is obtained by

$$f(x) \approx f(a) + f'(a)(x - a)$$

where $f'(a) = \frac{df(x)}{dx}|_{x=a}$. The expression on the right-hand side is just the equation for the tangent line to the graph $f(x)$ at point a. The above is true when x is close to a.



Example: Suppose $f(x) = x^2 - 5x + 7$ is to be linearised at $x = 5$.

$$f(5) = 25 - 25 + 7 = 7$$

$$f'(x) = 2x - 5$$

$$f'(5) = 10 - 5 = 5$$

Thus

$$f(x) \approx f(5) + (x - 5) * f(5)$$

$$= 7 + (x - 5) * 5$$

$$= 5x - 18$$

Hence can also say that $g(x) = 5x - 18$ is equation of tangent to curve $f(x) = x^2 - 5x + 7$ at $x = 5$.

Linear approximation of a multivariate function can be obtained similarly with the derivative at a point replaced by the partial derivatives. Given a differentiable scalar multivariate function $f(\mathbf{x})$, where $\mathbf{x} = [x_1, x_2]^T$, the linear approximation of the function at point $\mathbf{a} = [a_1, a_2]^T$ is

$$f(\mathbf{x}) \approx f(\mathbf{a}) + \frac{\partial}{\partial x_1} f(\mathbf{x})|_{\mathbf{x}=\mathbf{a}}(x_1 - a_1)$$

$$+\frac{\partial}{\partial x_2} f(\mathbf{x})|_{\mathbf{x}=\mathbf{a}}(x_2 - a_2)$$

$$= f(\mathbf{a}) + (\mathbf{x} - \mathbf{a})^T \nabla f(\mathbf{x})|_{\mathbf{x}=\mathbf{a}}$$

where

$$\nabla f(\mathbf{x}) = \left[\begin{array}{c} \frac{\partial}{\partial x_1} f(\mathbf{x}) \\ \frac{\partial}{\partial x_2} f(\mathbf{x}) \end{array} \right]$$

<u>Example:</u> Suppose

$$f(\mathbf{x}) = x_1^2 + 2x_2^2 - 5x_1 - 8x_2 + 7$$

is to be linearised at $\mathbf{x} = [5\ 4]^T$.

$$f([5\ \ 4]^T) = 25 + 2 \times 16 - 5 \times 5 - 8 \times 4 + 7 = 7$$

$$\frac{\partial}{\partial x_1} f(\mathbf{x}) = 2x_1 - 5$$

$$\frac{\partial}{\partial x_1} f(\mathbf{x})|_{\mathbf{x}=[5\ \ 4]^T} = 10 - 5 = 5$$

$$\frac{\partial}{\partial x_2} f(\mathbf{x}) = 4x_2 - 8$$

$$\frac{\partial}{\partial x_2} f(\mathbf{x})|_{\mathbf{x}=[5\ \ 4]^T} = 16 - 8 = 8$$

Thus

$$f(x) \approx 7 + (x_1 - 5) * 5 + (x_2 - 4) * 8$$

$$= 5x_1 + 8x_2 - 50$$

Hence can also say that $g(\mathbf{x}) = 5x_1 + 8x_2 - 50$ is equation of tangent to surface $f(\mathbf{x}) = x_1^2 + 2x_2^2 - 5x_1 - 8x_2 + 7$ at $x = [5\ \ \ 4]^T$.

Taylor series: Given a differentiable scalar function $f(x)$ of the real variable x, its Taylor series reads

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2!}f''(a)(x-a)^2 + \dots$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(a)(x-a)^k$$

where $f''(a) = \frac{d^2 f(x)}{dx^2}|_{x=a}$, $f^{(n)}(a) = \frac{d^n f(x)}{dx^n}|_{x=a}$.

- In engineering, the Taylor series of a function at a certain point up to a finite order n is used to approximate the function.

- If $n = 1$, this reduces to the linear approximation.

<ins>Example:</ins> let $f(x) = x^2 e^x$. Show that the Taylor expansion of $f(x)$ around 0, and up to the 4th order, is

$$f(x) = x^2 + x^3 + \frac{1}{2}x^4 + \ldots$$

$$f(0) = 0^2 e^0 = 0$$

$$f'(x) = 2xe^x + x^2 e^x = (2x + x^2)e^x$$

$$f'(0) = 0$$

$$f''(x) = (2 + 2x)e^x + (2x + x^2)e^x$$

$$= (2 + 4x + x^2)e^x$$

$$f''(0) = 2$$

$$f^{(3)}(x) = (4 + 2x)e^x + (2 + 4x + x^2)e^x$$

$$= (6 + 6x + x^2)e^x$$

$$f^{(3)}(0) = 6$$

$$f^{(4)}(x) = (6 + 2x)e^x + (6 + 6x + x^2)e^x$$

$$= (12 + 8x + x^2)e^x$$

$$f^{(4)}(0) = 12$$

Thus

$$f(x) = \frac{2}{2!}x^2 + \frac{6}{3!}x^3 + \frac{12}{4!}x^4 + \dots$$

$$= x^2 + x^3 + \frac{1}{2}x^4 + \dots$$

The Taylor series expansion can be generalized to the function of m variables as in the following.

Taylor expansion of multivariate functions:

Given a differentiable scalar multivariate function $f(\mathbf{x})$, where $\mathbf{x} = [x_1, ..., x_m]^T$, the Taylor series of the function at point $\mathbf{a} = [a_1, ..., a_m]^T$

$$f(\mathbf{x}) = f(\mathbf{a}) + \frac{\partial}{\partial x_1} f(\mathbf{x})|_{\mathbf{x}=\mathbf{a}}(x_1 - a_1) + ...$$

$$+ \frac{\partial}{\partial x_m} f(\mathbf{x})|_{\mathbf{x}=\mathbf{a}}(x_m - a_m)$$

$$+\frac{1}{2}\frac{\partial^2}{\partial x_1^2} f(\mathbf{x})|_{\mathbf{x}=\mathbf{a}}(x_1 - a_1)^2 + ...$$

$$+\frac{1}{2}\frac{\partial^2}{\partial x_m^2} f(\mathbf{x})|_{\mathbf{x}=\mathbf{a}}(x_m - a_m)^2 + ...$$

$$= f(\mathbf{a}) + (\mathbf{x} - \mathbf{a})^T \nabla f(\mathbf{x})|_{\mathbf{x}=\mathbf{a}}$$

$$+\frac{1}{2}(\mathbf{x} - \mathbf{a})^T \mathbf{H}(\mathbf{x} - \mathbf{a})^T + ...$$

where

$$\nabla f(\mathbf{x}) = \left[\begin{array}{c} \frac{\partial}{\partial x_1} f(\mathbf{x}) \\ \frac{\partial}{\partial x_2} f(\mathbf{x}) \end{array} \right]$$

$$H = \frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x}^2}|_{\mathbf{x}=\mathbf{a}}$$

Note that the second term is an inner product of a gradient of $f(\mathbf{x})$ with the vector $\mathbf{x} - \mathbf{a}$ and the third term is a quadratic form with the Hessian matrix $\frac{\partial^2 f}{\partial \mathbf{x}^2}$.

Example:let $f(\mathbf{x}) = (x_1^2+2x_2^2-5x_1-8x_2+7)e^{x_1}$. Show that the Taylor expansion of $f(\mathbf{x})$ around $\mathbf{0}$, and up to the 2nd order,is

$$f(\mathbf{x}) = 7 + 2x_1 - 8x_2 - \frac{1}{2}x_1^2 - 8x_1x_2 + 2x_2^2 + \ldots$$

$$f([0\ \ 0]^T) = 7$$

$$\frac{\partial}{\partial x_1}f(\mathbf{x}) = (2x_1-5)e^{x_1}+(x_1^2+2x_2^2-5x_1-8x_2+7)e^{x_1}$$

$$= (x_1^2 + 2x_2^2 - 3x_1 - 8x_2 + 2)e^{x_1}$$

$$\frac{\partial}{\partial x_1}f(\mathbf{x})|_{\mathbf{x}=[0\ \ 0]^T} = 2$$

$$\frac{\partial}{\partial x_2}f(\mathbf{x}) = (4x_2 - 8)e^{x_1}$$

$$\frac{\partial}{\partial x_2}f(\mathbf{x})|_{\mathbf{x}=[0\ \ 0]^T} = -8$$

$$\frac{\partial^2}{\partial x_1^2} f(\mathbf{x}) = (2x_1 - 3)e^{x_1} + (x_1^2 + 2x_2^2 - 3x_1 - 8x_2 + 2)e^{x_1}$$

$$= (x_1^2 + 2x_2^2 - x_1 - 8x_2 - 1)e^{x_1}$$

$$\frac{\partial^2}{\partial x_1^2} f(\mathbf{x})|_{\mathbf{x}=[0\ \ 0]^T} = -1$$

$$\frac{\partial^2}{\partial x_1 \partial x_2} f(\mathbf{x}) = (4x_2 - 8)e^{x_1}$$

$$\frac{\partial^2}{\partial x_1 \partial x_2} f(\mathbf{x})|_{\mathbf{x}=[0\ \ 0]^T} = -8$$

$$\frac{\partial^2}{\partial x_2 \partial x_1} f(\mathbf{x}) = (4x_2 - 8)e^{x_1}$$

$$\frac{\partial^2}{\partial x_2 \partial x_1} f(\mathbf{x})|_{\mathbf{x}=[0\ \ 0]^T} = -8$$

$$\frac{\partial^2}{\partial x_2^2} f(\mathbf{x}) = 4e^{x_1}$$

$$\frac{\partial^2}{\partial x_2^2} f(\mathbf{x})_{\mathbf{x}=[0\ \ 0]^T} = 4$$

Thus

$$f(x) \approx 7 + [x_1\ \ x_2] \times [2\ \ -8]^T$$

$$+\frac{1}{2}[x_1\ \ x_2] \times \begin{pmatrix} -1 & -8 \\ -8 & 4 \end{pmatrix} \times [x_1\ \ x_2]^T$$

Similarly for a scalar function of a matrix variable:

$$f(\mathbf{X}) = f(\mathbf{A}) + \mathtt{trace}[(\frac{\partial f}{\partial \mathbf{X}})^T(\mathbf{X} - \mathbf{A})]) + \ldots$$

The above formula shows 2 terms of Taylor expansion. It uses the extension of definition of an inner product to matrices

$$\mathbf{AB} = \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij} b_{ij}$$

but

$$\mathtt{trace}(\mathbf{A}^T\mathbf{B}) = \sum_{i=1}^{m} (\mathbf{A}^T\mathbf{B})_{ii} = \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij} b_{ij}$$