Introduction

Taylor series is a representation of a function as an infinite sum of terms that are calculated from the values of the function's derivatives at a single point. Any finite number of initial terms of the Taylor series of a function is called a Taylor polynomial. In single variable calculus, Taylor polynomial of n degrees is used to approximate an (n+1)-order differentiable function and the error of the approximation can be estimated by the (n+1)-th term of the Taylor series. By introducing vector and matrix calculus notations, we can express the same idea for multivariate functions and vector functions.

Applications

Because all numbers that can be represented by finite digits are rational numbers, the numerical computation of an irrational function at a particular point is almost always approximated. The first order and second order of Taylor polynomials are most frequently selected as the proper rational function to approximate irrational functions. This idea is called linear and quadratic approximation in calculus, respectively. In addition, the quadratic approximation is also used to in optimization because local maximum or minimum occurs at the critical points where the second term (first derivatives) of the Taylor polynomial is zero and the third term (second derivatives) are definitely positive or negative. In order to obtain the first or second order Taylor polynomial, we compute the coefficients of Taylor series by calculating the first and second derivatives of the original function. When we move towards the advanced mathematical applications (temperature in 4 dimensional temporal-spatial space and vector field of moving hurricane centers), we need to use multivariate (vector) functions, instead of single variable functions. In terms of the linear and quadratic approximation, we still use the idea of first and second order Taylor polynomials. However, we need to first generalize the concepts of the first and second order derivatives in multivariate context to obtain the coefficients of Taylor polynomials. Then, we can obtain the multivariate Taylor polynomial to approximate an irrational multivariate function.

Goal and Objectives

We will extend the concepts of the first and second derivatives in the context of multivariate functions and apply these concepts to obtain the first and second order Taylor polynomials for multivariate functions. Our objectives are to learn the following concepts and associative formulas:

1. Gradient vector and matrix calculus
2. Linear approximation multivariate functions
3. Quadratic Taylor formula for multivariate functions
4. Use MATLAB to compute the Taylor series

Reflection Questions

In history, mathematicians had to spend years calculating the value tables of many special functions such as Bessel functions and Legendre function. Nowadays, it is a trivial click to use MATLAB to estimate the value of any known function at any particular point. It seems that it is unnecessary to learn the approximating techniques. But, think about these questions:

1. How do you render a smooth surface to graph a multivariate function?
2. Why we only need to consider the first and second derivative to find optimal solutions to most applications?
3. What does an IMU (Inertial Measurement Unit) for robotic navigation system need to sense in order to estimate its own positions?
1. Gradient vector and matrix calculus

In single variable calculus, we study functions one variable scalar functions $y = f(x)$ that maps a scalar in $\mathbb{R}$ (real number) to another scalar in $\mathbb{R}$. The derivative stands for the slope of the tangent line to the curve defined by the functions.

In multivariate calculus, we study a multivariate scalar function $w = f(x, y, z)$ that maps a vector $\mathbf{v} = [x, y, z]'$ in $\mathbb{R}^3$ to a scalar in $\mathbb{R}$.

In advance mathematics courses for engineering and physics, we may have to use multivariate vector functions $\mathbf{v} = [x, y, z]'$ in $\mathbb{R}^3$ that maps $\mathbf{v} = [x, y, z]'$ in $\mathbb{R}^3$ to another vector $\mathbf{w} = [w_1, w_2, w_3]$ in $\mathbb{R}^3$.

Since the goal of the course is to model and simulate real world applications in 3D to 4D temporal-spatial space, we would like to use either multivariate scalar or vector functions. So, we need to express the same set of concepts such as derivative and second order derivative, Taylor formula, etc. by using a set of new notations, mostly in vector or matrix form.

For multivariate scalar valued function $w = f(x, y, z)$, $\frac{\partial w}{\partial x}$, $\frac{\partial w}{\partial y}$, and $\frac{\partial w}{\partial z}$ are three partial derivatives of the function $f(x, y, z)$ with respect to the three variable $x$, $y$, and $z$. Sometime, we use the short hand notations $f_x$, $f_y$, and $f_z$ to express the three partial derivatives.

The gradient vector of $w = f(x, y, z)$

$$\nabla f = \left[ \frac{\partial w}{\partial x}, \frac{\partial w}{\partial y}, \frac{\partial w}{\partial z} \right] = [f_x, f_y, f_z] \quad (1.1)$$

is analogically equivalent to the derivative of single variable functions. For examples, the critical points of function $y = f(x)$ are the points where the derivative $f'(x) = 0$ or does not exist. Similarly, the critical points of multivariate function $w = f(x, y, z)$, are the points where the gradient $\nabla f = [0,0,0]$ or does not exist. The gradient of a function describes the orientation of a tangent plane to the level surface defined by implicit function $f(x, y, z) = 0$ at a given point.

Example 1

Given $w = f(x, y, z) = e^{2x}y^3 + \sin(z)$ and the level surface $f(x, y, z) = 1$ at a point $P(0,1,\pi)$, find the gradient function, and general equation of the tangent plane to the level surface.

Solution:

Calculate the three partial derivatives of $f$, we have

$$f_x = 2e^{2x}y^3, f_y = 3e^{2x}y^2, f_z = \cos(z),$$
Therefore, the gradient is:

$$\nabla f = [2e^{2x}y^3, 3e^{2x}y^2, \cos(z)]^T.$$ 

Substitute the component values of $P$ to the gradient, we have $\nabla f|_P = [2, 3, -1]$, which is the normal vector of the tangent plane to the level surface. Hence, the point normal equation of the tangent plane is:

$$2(x - 0) + 3(y - 1) + (-1)(z - \pi) = 0.$$ 

The general equation is

$$2x + 3y - z = 3 - \pi.$$ 

**Component-wise differentiation**

For multivariate vector valued functions $\vec{W} = [f_1(x, y, z), f_2(x, y, z), f_3(x, y, z)]$ that maps $\vec{v} = [x, y, z]'$ in $\mathbb{R}^3$ to another vector $\vec{W} = [w_1, w_2, w_3]$ in $\mathbb{R}^3$, we can define the so called Jacobian Matrix $J(\vec{W})$ and Jacobian determinant as follows:

$$J(\vec{W}) = \begin{bmatrix}
\frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\
\frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \\
\frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z}
\end{bmatrix} \quad \text{and} \quad |J(\vec{W})| = \begin{vmatrix}
\frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\
\frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \\
\frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z}
\end{vmatrix} \quad (1.2)$$

Notice that each row of the 3X3 matrix is the gradient of each component scalar valued function and each column is the partial derivatives of the vector function with respect to an individual variable. The Jacobian of a function describes the orientation of a tangent plane to the function at a given point. In this way, the Jacobian generalizes the gradient of a scalar valued function of multiple variables which itself generalizes the derivative of a scalar-valued function of a scalar.

**Example 2**

Notice that gradient in example 1 above is a vector function, find the Jacobian matrix as a matrix function and Jacobian determinant of the gradient at point $P(0, 1, \pi/2)$.

$$\vec{w} = \nabla f = [2e^{2x}y^3, 3e^{2x}y^2, \cos(z)].$$

Solution: For vector valued function $\vec{w}$, $f_1 = 2e^{2x}y^3, f_2 = 3e^{2x}y^2, f_3 = \cos(z)$
Multivariate Approximation and Matrix Calculus

\[ f(\mathbf{w}) = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} \end{bmatrix} = \begin{bmatrix} 4e^{2x}y^3 & 6e^{2x}y^2 & 0 \\ 6e^{2x}y^2 & 6e^{2x}y & 0 \\ 0 & 0 & -\sin(z) \end{bmatrix} \]

\[ |J(\mathbf{w})| = \begin{vmatrix} 4e^{2x}y^3 & 6e^{2x}y^2 & 0 \\ 6e^{2x}y^2 & 6e^{2x}y & 0 \\ 0 & 0 & -\sin(z) \end{vmatrix}, \]

\[ |J(\mathbf{w})|_p = \begin{vmatrix} 4 & 6 & 0 \\ 6 & 6 & 0 \\ 0 & 0 & -1 \end{vmatrix} = -1(24 - 36) = 12 \]

**Second Derivative**

The second derivative of a multivariate scalar valued function \( w = f(x, y, z) \) is the Hessian Matrix, which is defined as

\[ H(w) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z^2} \end{bmatrix} = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix} \quad (1.3) \]

**Example 3**

Find the Hessian Matrix of the scalar-valued multivariate function \( w = f(x, y, z) = e^{2x}y^3 + \sin(z) \).

Solution:

\[ f_x = 2e^{2x}y^3, f_y = 3e^{2x}y^2, f_z = \cos(z) \]
\[ f_{xx} = 4e^{2x}y^3, f_{xy} = 6e^{2x}y^2, f_{xz} = 0 \]
\[ f_{yx} = 6e^{2x}y^2, f_{yy} = 6e^{2x}y, f_{yz} = 0 \]
\[ f_{zx} = 0, f_{zy} = 0, f_{zz} = -\sin(z) \]

Hence the Hessian Matrix is:

\[ H(w) = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix} = \begin{bmatrix} 4e^{2x}y^3 & 6e^{2x}y^2 & 0 \\ 6e^{2x}y^2 & 6e^{2x}y & 0 \\ 0 & 0 & -\sin(z) \end{bmatrix} \]

Compare the result of the second and third examples, we find that the Hessian Matrix is actually the Jacobian Matrix of the gradient vector function. This is true for any second differentiable multivariate functions. So, other authors may use \( \nabla^2 f = |(\nabla f)| \) to express the Hessian Matrix.

For a more general \( n \)-variable scalar-valued function \( w = f(x_1, x_2, \ldots, x_n) \) that maps a vector \([x_1, x_2, \ldots, x_n]\) in \( R^n \) to a scalar in \( R \), we can write the gradient of \( f \) as

\[ \nabla f = \left[ \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right] = \left[ f_{x_1}, f_{x_2}, \ldots, f_{x_n} \right] \text{ in } R^n \quad (1.4) \]
The Hessian Matrix as the following $n \times n$ matrix

$$H(w) = \begin{bmatrix}
\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\
\frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2}
\end{bmatrix} \tag{1.5}$$

**Self-Check Exercises**

1. Find the gradient and Hessian Matrix of the function $w = f(x, y, z) = ye^{2x} + xy^3 + x^2\sin(z)$ at point $P(0,1,\pi/2)$.

2. Find the Jacobian matrix and the Jacobian determinant of the vector function $\vec{w} = [2 ye^{2x} + y^3 + 2x\sin(z), e^{2x} + 3xy^2, x^2\cos(z)]$ at point $P(0,1,\pi/2)$. 
2. Linear approximation for multivariate functions

**Linear approximation**

A differentiable function \( f(x) \) is one that resembles a linear function (straight line) at close range. The linear approximation to \( f \) at a point \( x_0 \) is the linear function it resembles there. If \( f(x) \) has second derivatives at a point \( x_0 \), the quadratic function having the same second derivatives is the quadratic approximation to it at the point \( x_0 \).

In one dimension the graph of a function against the independent variable is a curve, and the linear approximation to it is the function whose graph is the *tangent line* to it at \( x_0 \).

In two dimensions, the graph of \( f(x, y) \) against the two independent variables, is a surface, and the linear approximation to it at \( P_0 \) is the *plane tangent to that surface at \( P_0 \).*

**In single variable form**

In *one dimension* the linear function \( L \), \( L = a(x - x_0) + b \) is determined by the conditions that

\[
L(x_0) = f(x_0) \text{ and } L'(x_0) = f'(x_0)
\]

We get

\[
a = f'(x_0), \quad b = f(x_0)
\]

For three variable function \( f(x, y, z) \), if we write the linear approximation function of \( f \) at \( P_0(x_0, y_0, z_0) \) as \( L \):

\[
L(x, y, z) = a_1(x-x_0) + a_2(y-y_0) + a_3(z-z_0) + b = \ddot{a} \cdot \ddot{x} + b
\]

we obtain

\[
\ddot{a} = \nabla f(P_0), \quad b = f(P_0)
\]

The obvious use of the linear approximation is in estimating the value of a function at \( P_1(x_1, y_1, z_1) \) knowing its value at \( P_0(x_0, y_0, z_0) \) and its gradient there.

**In multivariate form**

\[
L(P_1) = \nabla f(P_0) \cdot \overline{P_0P_1} + f(P_0), \quad \text{where } \overline{P_0P_1} = [x_1-x_0, y_1-y_0, z_1-z_0]' \quad (2.1)
\]

When point \( P_1(x_1, x_2, ... x_n) \) is near to the point \( P_0(x_{10}, x_{20}, ... x_{n0}) \) a multivariate function, \( f(x_1, x_2, ... x_n) \) at \( P_1 \) can be approximated by the linear functions

\[
L(P_1) = \nabla f(P_0) \cdot \overline{P_0P_1} + f(P_0), \quad \text{where } \overline{P_0P_1} = [x_1-x_{10}, x_2-x_{20}, ... x_n-x_{n0}]'
\]

When point \( P_1(x_1, x_2, ... x_n) \) is near to the point \( P_0(x_{10}, x_{20}, ... x_{n0}) \) a multivariate function, a vector function with \( n \) components

\[
\vec{F} = [f_1(x_1, x_2, ... x_n), f_2(x_1, x_2, ... x_n), ..., f_m(x_1, x_2, ... x_n)]' \quad \text{at } P_1 \text{ can be approximated by the linear vector functions}
\]
\[ \mathbf{L}(P_1) = J(\mathbf{F}(P_0)) \mathbf{P}_0 \mathbf{P}_1^T + \mathbf{F}(P_0), \quad \text{where} \quad \mathbf{P}_0 \mathbf{P}_1^T = [x_1 - x_{10}, \ x_2 - x_{20}, \ldots, x_n - x_{n0}]^T, \]

Where \( J(\mathbf{F}(P_0)) \) is the \( m \times n \) Jacobian Matrix of the function \( \mathbf{F} \) at point \( P_0 \).

We can simply regard the last formula as a list of \( m \) previous formulas.

Use the linear approximation of \( f(x, y, z) = \sqrt[3]{x \sqrt[4]{y \sqrt[3]{z}}} \) at point \( P_0(27, 25, 16) \) to estimate \( \sqrt[3]{28 \sqrt[4]{24 \sqrt[3]{15}}} \).

Solution: The value of the function \( f \) at \( P_1(28, 24, 15) \) can be approximated by the value of \( L \) at \( P_1(28, 24, 15) \), where the linear function \( L \) is given by eq. (1.2).

\[
\nabla f(P_0) = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{bmatrix}(P_0) = \begin{bmatrix} \frac{1}{3} x^{-2/3} y^{1/2} z^{1/4} & \frac{1}{2} x^{1/3} y^{-1/2} z^{1/4} & \frac{1}{4} x^{1/3} y^{1/2} z^{-3/4} \end{bmatrix}(P_0),
\]

\[
= \begin{bmatrix} \frac{1}{2} 27^{-2/3} 25^{1/2} 16^{1/4} & \frac{1}{2} 27^{1/3} 25^{-1/2} 16^{1/4} & \frac{1}{4} 27^{1/3} 25^{1/2} 16^{-3/4} \end{bmatrix} = \begin{bmatrix} 10 & 3 & 15 \end{bmatrix},
\]

\[
\mathbf{P}_0 \mathbf{P}_1 = [x_1 - x_{10}, \ y_1 - y_{10}, \ z_1 - z_{10}]^T = [28 - 27, 24 - 25, 15 - 16] = [1, -1, -1] ,
\]

\[
f(P_0) = 30, L(P_1) = \nabla f(P_0) \cdot \mathbf{P}_0 \mathbf{P}_1^T + f(P_0) = \begin{bmatrix} 10 & 3 & 15 \end{bmatrix} [1, -1, -1]^T + 30,
\]

\[
= \frac{10}{27} - \frac{3}{5} - \frac{15}{32} + 30 = 29 \frac{1303}{4320}
\]

Use Maple, you will find the error is

\[
29 \frac{1303}{4320} - \sqrt[3]{24.0} \cdot (15.0) \cdot \frac{1}{3} \cdot (28.0)
\]

\[
= 29.30162037 - 29.27618252 = 0.02543785.
\]

**Self-Check Exercises**

(3) Use the linear approximation of \( f(x, y) = \sqrt[3]{x} \sqrt[4]{y} \) at point \( P_0(27, 25) \) to estimate \( \sqrt[3]{27.5 \sqrt[4]{24.5}} \).

(4) Use the linear approximation of \( f(x, y, z) = (z + 1)e^{x+2y} \) at point \( P_0(0, 0, 0) \) to estimate \( f(0.1, 0.1, 0.2) \).
3. Quadratic approximation to multivariate functions

Concept

In one dimension, the quadratic function $Q(x)$ to approximate the second order differentiable function $f(x)$ is,

$$Q(x) = a + b(x - x_0) + c(x - x_0)^2,$$  \hspace{1cm} (3.1)

Generalization

which is determined by the conditions that

$$Q(x_0) = f(x_0) \text{ and } Q'(x_0) = f'(x_0), \text{and } Q''(x_0) = f''(x_0).$$

We get

$$a = f(x_0), \ b = f'(x_0), c = \frac{1}{2} f''(x_0)$$

$Q(x)$ is actually the second degree Taylor polynomial approximation of $f(x)$.

Quadratic Approximation

In two dimensions, the quadratic approximation which we write out in detail, is of great use in determining the nature of a critical point at $P_0\equiv(x_0,y_0)$, and can be useful in approximating $f(P)$ at $P(x,y)$ when the linear approximation is insufficiently accurate. We denote the quadratic approximating function to $f$ at $P_0$ by $Q$.

In quadratic Form

$$Q(P) = f(P_0) + \nabla f(P_0) \cdot \overrightarrow{P_0P} + \frac{1}{2} \left[ \frac{\partial^2 f}{\partial x^2} (x-x_0)^2 + (x-x_0)(y-y_0) \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y^2} (y-y_0)^2 \right] \bigg|_{P_0}$$

$$= f(P_0) + \nabla f(P_0) \cdot \overrightarrow{P_0P} + \frac{1}{2} \overrightarrow{P_0P} H(f) \overrightarrow{P_0P}'$$

$$= f(P_0) + \nabla f(P_0) \cdot \overrightarrow{P_0P} + \frac{1}{2} \overrightarrow{P_0P} H(f) \overrightarrow{P_0P}'$$

Where $H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$ is the Hessian Matrix $f(P)$.  \hspace{1cm} (3.2)

Find the second-order Taylor polynomial for $f(x,y) = e^{-2x+y}$ at $P_0(0,0)$ and use it to estimate the value of the function at $R(0.1,0.1)$

Solution:

$$f(0,0) = 1, \ \nabla f = [-2 e^{-2x+y}, e^{-2x+y}], \ \nabla f|_{P_0} = [-2,1],$$

$$H(f) = |(\nabla f)| = \begin{bmatrix} 4e^{-2x+y} & -2e^{-2x+y} \\ -2e^{-2x+y} & e^{-2x+y} \end{bmatrix}, \ H(f)|_{P_0} = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}$$

The second order Taylor polynomial at point $P(x,y)$ is

$$Q(P) = f(P_0) + \nabla f(P_0) \cdot \overrightarrow{P_0P} + \frac{1}{2} \overrightarrow{P_0P} H(f) \overrightarrow{P_0P}'$$

$$= 1 + [-2,1] \cdot [x,y] + \frac{1}{2} \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= 1 - 2x + y + 2x^2 - 2xy + \frac{1}{2} y^2$$
Multivariate Approximation and Matrix Calculus

The estimate is: \( f(R) \approx Q(R) = 1 - 0.2 + 0.1 + 0.02 - 0.02 + 0.005 = 0.905 \).

It is easy to check that functions \( f \) and \( Q \) share the same value, same gradient and same second order derivatives at the point \( P_0 \).

For a more general \( n \)-variable scalar-valued function \( w = f(P) \) that maps a vector \( P = [x_1, x_2, ..., x_n]^T \in \mathbb{R}^n \) to a scalar in \( \mathbb{R} \), we can write the quadratic approximation of the function \( w = f(P) \) near point \( P_0 = [c_1, c_2, ..., c_n]^T \in \mathbb{R}^n \) by

\[
\begin{align*}
f(P) & \approx Q(P) = f(P_0) + \nabla f(P_0) \cdot \bar{P}_0 \bar{P}^T + \frac{1}{2} \bar{P}_0 \bar{P}^T H(f) \bar{P}_0 \bar{P}^T \\
& \text{where } \bar{P}_0 \bar{P} = [x_1 - c_1, x_2 - c_2, ..., x_n - c_n]^T \in \mathbb{R}^n,
\end{align*}
\]

(3.3)

Where the gradient and the Hessian matrix are defined in equations (1.4) and (1.5) in the first section.

**Example 6**

Find both the linear and quadratic approximation for \( f(x, y, z) = e^{-2x+y+3z} \) at \( P_0(0,0,0) \) and estimate the value of the function at \( R(-0.1,0.1,0.1) \) by both linear and quadratic approximations.

Solution: \( \nabla f = [-2e^{-2x+y+3z}, e^{-2x+y+3z}, 3e^{-2x+y+3z}] \), \( \nabla f|_{P_0} = [-2,1,3] \),

\[
\begin{align*}
f(0,0,0) & = 1, \quad f(P) \approx L(P) = f(P_0) + \nabla f(P_0) \cdot \bar{P}_0 \bar{P}^T \\
& = 1 + [-2,1,3] \cdot [\begin{array}{c} x \\ y \\ z \end{array}] = 1 - 2x + y + 3z
\end{align*}
\]

\[
\begin{align*}
H(f) & = \nabla^2 f = \begin{bmatrix} 4e^{-2x+y+3z} & -2e^{-2x+y+3z} & -6e^{-2x+y+3z} \\
-2e^{-2x+y+3z} & e^{-2x+y+3z} & 3e^{-2x+y+3z} \\
-6e^{-2x+y+3z} & 3e^{-2x+y+3z} & 9e^{-2x+y+3z} \end{bmatrix} \\
H(f)|_{P_0} & = \nabla^2 f|_{P_0} = \begin{bmatrix} 4 & -2 & -6 \\
-2 & 1 & 3 \\
-6 & 3 & 9 \end{bmatrix}, \quad \bar{P}_0 \bar{P} = [x, y, z]
\end{align*}
\]

Notice that the Hessian matrix is a symmetric matrix. By the formula (3.2)

\[
\begin{align*}
f(P) & \approx Q(P) = f(P_0) + \nabla f(P_0) \cdot \bar{P}_0 \bar{P}^T + \frac{1}{2} \bar{P}_0 \bar{P}^T H(f) \bar{P}_0 \bar{P}^T \\
& = 1 + [-2,1,3] \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 4 & -2 & -6 \\
-2 & 1 & 3 \\
-6 & 3 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \\
& = 1 - 2x + y + 3z + 2x^2 - 2xy - 6xz + \frac{1}{2} y^2 + 3yz + \frac{9}{2} z^2
\end{align*}
\]

The linear approximation is \( f(R) \approx L(R) = 1 + 0.2 + 0.1 + 0.3 = 1.6 \).

The quadratic approximation is

\[
f(R) \approx Q(R) = 1 + 0.2 + 0.1 + 0.3 + 0.02 + 0.02 + 0.06 + 0.005 + 0.03 + 0.045 = 1.78
\]

While \( f(P) = e^{2(0.1)+0.1+3\cdot0.1} = e^{0.6} \approx 1.822118800 \), the quadratic approximate 1.78 is a much better estimate than the linear approximate 1.6.
Self-Check Exercises

5. (a) Use the quadratic approximation of \( f(x, y) = \sqrt{x} \sqrt{y} \) at point \( P_5(27, 25) \) to estimate \( \sqrt{27.5} \sqrt{24.5} \).

(b) Use your calculator to calculate the "accurate" answer and compare the relative accuracy of the approximation, where the relative accuracy is defined as the ratio of the error against the accurate answer.

4. MATLAB commands

Declare symbolic variables

\[ f = 12 + (x-1)(x-1)(x-2)(x-3); \]

Taylor Approximation in single variable form in different orders

\[ t1 = \text{taylor}(f, 1, 2.5); \quad \% \text{expansion with first term} \]
\[ t2 = \text{taylor}(f, 2, 2.5); \quad \% \text{linear expansion} \]
\[ t3 = \text{taylor}(f, 3, 2.5); \quad \% \text{quadratic expansion} \]
\[ t6 = \text{taylor}(f, 6, 2.5); \quad \% \text{fifth order expansion} \]

Plot graphs

\[ xd = 0:.05:4; \]
\[ fval = \text{double} \left( \text{subs}(f, x, xd) \right); \quad \% \text{original function} \]
\[ c1 = \text{plot}(xd, fval); \]
\[ \text{set} (c1, 'LineWidth', 2, 'Color', 'r', 'LineStyle', '-'); \]
\[ \text{ezplot}([t1, t2, t3, t4], [0, 4]); \quad \% \text{plot} \]
Use build in functions to define functions

Take derivatives in different orders

Define multivariate functions

Differentiation of a symbolic expression is performed by means of the function `diff`. For instance, let’s find the derivative of $f(x) = \sin(e^x)$.

```matlab
>> syms x
>> f=sin(exp(x))
```

For instance, let’s find the derivative of $f(x) = \sin(e^x)$.

```matlab
>> syms x
>> f=sin(exp(x))
```

```matlab
f =
  sin(exp(x))
>> diff(f)
ans =
  cos(exp(x))*exp(x)
```

Partial derivatives

In first order

The $n^{th}$ derivative of $f$ is $\text{diff}(f,n)$.

```matlab
>> diff(f,2)
ans =
-sin(exp(x))*exp(x)^2+cos(exp(x))*exp(x)
```

Third order

To compute the partial derivative of an expression with respect to some variable we specify that variable as an additional argument in `diff`. Let $f(x,y) = x^3y^4 + y\sin x$.

```matlab
>> syms x y
>> f=x^3*y^4+y*sin(x)
f =
x^3*y^4+y*sin(x)
>> diff(f,x)
ans =
3*x^2*y^4+y*cos(x)
```

Define a vector multivariate function

Next we compute $\frac{df}{dy}$.

```matlab
>> diff(f,y)
ans =
4*x^3*y^3+sin(x)
```

Jacobian matrix

Finally we compute $\frac{d^3 f}{dx^3}$.

```matlab
>> diff(f,x,3)
ans =
6*y^4-y*cos(x)
```
The Jacobian matrix of a function $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ can be found directly using the jacobian function. For example, let $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ be defined by $f(x,y)=(\sin(xy), x^2+y^2, 3x-2y)$.

```matlab
>> f=[sin(x*y); x^2+y^2; 3*x-2*y]
f =
[ sin(y*x)]
[ x^2+y^2]
[ 3*x-2*y]
>> Jf=jacobian(f)
Jf =
[ cos(y*x)*y, cos(y*x)*x]
[ 2*x, 2*y]
[ 3, -2]
```

In the case of a linear transformation, the Jacobian is quite simple.

```matlab
>> A=[11 -3 14 7;5 7 9 2;8 12 -6 3]
A =
11 -3 14 7
5 7 9 2
8 12 -6 3
>> syms x1 x2 x3 x4
>> x=[x1;x2;x3;x4]
x =
[ x1]
[ x2]
[ x3]
[ x4]
>> T=A*x
T =
[ 11*x1-3*x2+14*x3+7*x4]
[ 5*x1+7*x2+9*x3+2*x4]
[ 8*x1+12*x2-6*x3+3*x4]
```

Now let’s find the Jacobian of $T$.

```matlab
>> JT=jacobian(T)
JT =
[ 11, -3, 14, 7]
[ 5, 7, 9, 2]
[ 8, 12, -6, 3]
```

The Jacobian of $T$ is precisely $A$.

Next suppose $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is a scalar valued function. Then its Jacobian is just its gradient. (Well, almost. Strictly speaking, they are the transpose of one another since the Jacobian is a row vector and the gradient is a column vector.) For example, let $f(x,y)=(4x^2-1)e^{x^2+y^2}$.

```matlab
>> syms x y real
>> f=(4*x^2-1)*exp(-x^2-y^2)
f =
(4*x^2-1)*exp(-x^2-y^2)
>> gradf=jacobian(f)
gradf =
[ 8*x*exp(-x^2-y^2)-2*(4*x^2-1)*x*exp(-x^2-y^2), -2*(4*x^2-1)*y*exp(-x^2-y^2)]
```

Next we use solve to find the critical points of $f$. 
Multivariate Approximation and Matrix Calculus

$$\begin{align*}
\text{S} &= \text{solve}([\text{gradf}(1), \text{gradf}(2)]) \\
\text{S} &= \begin{bmatrix}
0 \\
0 \\
\frac{1}{2}5^{1/2} \\
0 \\
\frac{-1}{2}5^{1/2} \\
0
\end{bmatrix}
\end{align*}$$

Thus the critical points are (0,0), ($\frac{\sqrt{5}}{2}$,0) and ($\frac{-\sqrt{5}}{2}$,0).

The Hessian of a scalar valued function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is the $n \times n$ matrix of second order partial derivatives of $f$. In MATLAB we can obtain the Hessian of $f$ by computing the Jacobian of the Jacobian of $f$. Consider once again the function $f(x,y)=(4x^2-1)e^{-x^2-y^2}$.

$$\begin{align*}
\text{x} &= \text{sym}(x) \\
\text{y} &= \text{sym}(y) \\
\text{Hf} &= \text{simple}(\text{hessian}(\text{hessian}(f))) \\
\text{Hf} &= \begin{bmatrix}
2\exp(-x^2-y^2)*(2*x^2-5), & 0 \\
4*x^2*y^2*\exp(-x^2-y^2)*(1+2*y^2), & 2\exp(-x^2-y^2)*(1+2*y^2)*(-1+2*y^2)
\end{bmatrix}
\end{align*}$$

We can now use the Second Derivative Test to determine the type of each critical point of $f$ found above.

$$\begin{align*}
\text{sub(Hf,}\{x,y\},\{0,0\}) &= \begin{bmatrix}
10 \\
0
\end{bmatrix} \\
\text{sub(Hf,}\{x,y\},\{1/2*5^{1/2},0\}) &= \begin{bmatrix}
0.0000 \\
2.9200
\end{bmatrix} \\
\text{sub(Hf,}\{x,y\},\{-1/2*5^{1/2},0\}) &= \begin{bmatrix}
0.0000 \\
2.9200
\end{bmatrix}
\end{align*}$$

Thus $f$ has a local minimum at (0,0) and local maxima at the other two critical points. Evaluating $f$ at the critical points gives the maximum and minimum values of $f$.

$$\begin{align*}
\text{sub(f,}\{x,y\},\{0,0\}) &= -1 \\
\text{sub(f,}\{x,y\},\{1/2*5^{1/2},0\}) &= 4\exp(-5/4) \\
\text{sub(f,}\{x,y\},\{-1/2*5^{1/2},0\}) &= 4\exp(-5/4)
\end{align*}$$

Thus the minimum value of $f$ is $f(0,0)=-1$ and the maximum value is $f(\sqrt{5}/2,0)=f(-\sqrt{5}/2,0)=4e^{-5/4}$.

The graph of $f$ is shown in figure 1.
1. Use your scientific calculator to find the answers
   (a) Find the linear approximation of \( f(x, y, z) = ye^{2x} + xy^3 + x^2 \sin(x) \)
       at point \( P(0.1, \pi/2) \), use what you find to estimate \( f(0.1, 0.98, 3.2/2) \).
   (b) Find the quadratic approximation of \( f(x, y, z) = ye^{2x} + xy^3 + x^2 \sin(x) \)
       at point \( P(0, 1, \pi/2) \) and use what you find estimate \( f(0.1, 0.98, 3.2/2) \).
   (b) Use your calculator to calculate the “accurate” answer and compare
       the relative accuracy of the two approximations above, where the
       relative accuracy is defined as the ratio of the error against the
       accurate answer.

2. Use MATLAB command to find the answers to the problem 1.
Answers to Self-Check Exercises

1. Answer \( \nabla w = \left[ 2 ye^{2x+y^3} + 2x \sin(z), e^{2x+3xy^2}, x^2 \cos(z) \right] \left( \frac{0}{2}, \frac{\pi}{2} \right) = [3, 10] \)

2. Answer, the Jacobian matrix is the matrix consisting of the three gradient of the vector function above

\[
\begin{bmatrix}
  4 ye^{2x} + 2 \sin(z) & 2 e^{2x+3y^2} & 2 \cos(z) \\
  2 e^{2x+3y^2} & 6xy & 0 \\
  2x \cos(z) & 0 & -x^2 \sin(z)
\end{bmatrix}
\left( \frac{0}{2}, \frac{\pi}{2} \right) =
\begin{bmatrix}
  6 & 5 & 0 \\
  5 & 0 & 0 \\
  0 & 0 & 0
\end{bmatrix}
\]

3. Let \((x_0, y_0) = (27, 25), (\Delta x, \Delta y) = (0.5, 0.5), f(27, 25) = 3 \times 5 = 15, f_x(27, 25) = \frac{1}{3} \sqrt{y}/\sqrt{x^2} = 5/27, f_y(27, 25) = \frac{1}{2} \sqrt{x}/\sqrt{y} = 3/10, \nabla f(17, 25) = [\frac{5}{27}, \frac{1}{10}] \cdot (\Delta x, \Delta y) = 15 + \frac{5}{27} \cdot \frac{3}{10} = 15.2426 \)

4. \((\Delta x, \Delta y, \Delta z) = (0.1, 0.1, 0.2), f(0, 0, 0) = 1, \nabla f(0, 0, 0) = [e^{2x+y}, 2(z + 1)e^{2x+y}, (z + 1)e^{2x+y}](0, 0, 0) = [1, 2, 1], f(0, 1, 0, 1, 0.2) \approx f(0, 0, 0) + \nabla f(0, 0, 0) \cdot [0.1, 0.1, 0.2] = 1 + 0.5 = 1.5 \)

5. Let \((x_0, y_0) = (27, 25), (\Delta x, \Delta y) = (0.5, -0.5), f(27, 25) = 3 \times 5 = 15, f_x(27, 25) = \frac{1}{3} \sqrt{y}/\sqrt{x^2} = 5/27, f_y(27, 25) = \frac{1}{2} \sqrt{x}/\sqrt{y} = 3/10, \nabla f(17, 25) = [\frac{5}{27}, \frac{1}{10}] \cdot (\Delta x, \Delta y) = 15 + \frac{5}{27} \cdot \frac{3}{10} = 15.2426 \)

\[
L((27.5, 24.5)) = f(27.25) + \nabla f(27.25) \cdot [0.5, -0.5]' = 15 + \frac{5}{27} \cdot \frac{3}{10} \cdot [0.5, -0.5]' = 14.9690
\]

\[
f_x(27.25) = \frac{1}{3} \sqrt{y}/\sqrt{x^2} = \frac{10}{37}, f_y(27.25) = \frac{1}{2} \sqrt{x}/\sqrt{y} = \frac{1}{3}, f_{xy}(27.25) = \frac{1}{10} \sqrt{x}/\sqrt{y} = \frac{1}{90}
\]

\[
H(f) =
\begin{bmatrix}
  \frac{10}{37} & 1 \\
  1 & \frac{1}{90}
\end{bmatrix}
\]

\[
Q(27.5, 24.5) = f(27.25) + \nabla f(27.25) \cdot [0.5, -0.5]' + \frac{1}{2} [0.5, -0.5] H(f) [0.5, -0.5]'
\]

\[
= 15 + \frac{5}{27} \cdot \frac{3}{10} + \frac{10}{37} \cdot [0.5, -0.5] + \frac{1}{2} [0.5, -0.5] \begin{bmatrix}
  \frac{10}{37} & 1 \\
  1 & \frac{1}{90}
\end{bmatrix} [0.5, -0.5]
\]

\[
= 15 + \left( \frac{5}{27} \cdot \frac{3}{10} - \frac{10}{37} \cdot \frac{1}{4} \right) + \left( \frac{1}{2} - \frac{3}{500} \right) \cdot [0.5, -0.5] = 14.9649
\]

The "accurate" answer from MATLAB is 14.9403.
The absolute error for linear approximation is 14.9690-14.9403 = 0.0287,
Relative error is 0.19% 
The absolute error for quadratic approximation is =14.9649-14.9403=0.0246
Relative error is 0.16%.

Remark: In this example, the quadratic approximation is only slightly better than the linear approximation. Try both estimates for \(f(27.5, 25.5)\) and discuss the differences.