Section 4.2

Best Affine Approximations

Best affine approximations

The following definitions should look very familiar.

Definition Suppose $f : \mathbb{R}^m \to \mathbb{R}^n$ is defined on an open ball containing the point $\mathbf{c}$. We call an affine function $A : \mathbb{R}^m \to \mathbb{R}^n$ the *best affine approximation* to f at $\mathbf{c}$ if (1) $A(\mathbf{c}) = f(\mathbf{c})$ and (2) $\|R(\mathbf{h})\|$ is $o(\mathbf{h})$, where

$$R(\mathbf{h}) = f(\mathbf{c}+\mathbf{h}) - A(\mathbf{c}+\mathbf{h}). \tag{4.2.1}$$

Suppose $A : \mathbb{R}^n \to \mathbb{R}^n$ is the best affine approximation to f at $\mathbf{c}$. Then, from our work in Section 1.5, there exists an $n \times m$ matrix M and a vector $\mathbf{b}$ in $\mathbb{R}^n$ such that

$$A(\mathbf{x}) = M\mathbf{x} + \mathbf{b} \tag{4.2.2}$$

for all $\mathbf{x}$ in $\mathbb{R}^m$. Moreover, the condition $A(\mathbf{c}) = f(\mathbf{c})$ implies $f(\mathbf{c}) = M\mathbf{c} + \mathbf{b}$, and so $\mathbf{b} = f(\mathbf{c}) - M\mathbf{c}$. Hence we have

$$A(\mathbf{x}) = M\mathbf{x} + f(\mathbf{c}) - M\mathbf{c} = M(\mathbf{x} - \mathbf{c}) + f(\mathbf{c}) \tag{4.2.3}$$

for all $\mathbf{x}$ in $\mathbb{R}^m$. Thus to find the best affine approximation we need only identify the matrix M in (4.2.3).

Definition Suppose $f : \mathbb{R}^m \to \mathbb{R}^n$ is defined on an open ball containing the point $\mathbf{c}$. If f has a best affine approximation at $\mathbf{c}$, then we say f is *differentiable* at $\mathbf{c}$. Moreover, if the best affine approximation to f at $\mathbf{c}$ is given by

$$A(\mathbf{x}) = M(\mathbf{x} - \mathbf{c}) + f(\mathbf{c}), \tag{4.2.4}$$

then we call M the *derivative* of f at $\mathbf{c}$ and write $Df(\mathbf{c}) = M$.

Now suppose $f : \mathbb{R}^m \to \mathbb{R}^n$ and A is an affine function with $A(\mathbf{c}) = f(\mathbf{c})$. Let f_k and A_k be the kth coordinate functions of f and A, respectively, for $k = 1, 2, \ldots, n$, and let R be the remainder function

$$\begin{aligned} R(\mathbf{h}) &= f(\mathbf{c}+\mathbf{h}) - A(\mathbf{c}+\mathbf{h}) \\ &= (f_1(\mathbf{c}+\mathbf{h}) - A_1(\mathbf{c}+\mathbf{h}), f_2(\mathbf{c}+\mathbf{h}) - A_2(\mathbf{c}+\mathbf{h}), \ldots, f_n(\mathbf{c}+\mathbf{h}) - A_n(\mathbf{c}+\mathbf{h})). \end{aligned}$$

Then

$$\frac{R(\mathbf{h})}{\|\mathbf{h}\|} = \left(\frac{f_1(\mathbf{c}+\mathbf{h}) - A_1(\mathbf{c}+\mathbf{h})}{\|\mathbf{h}\|}, \frac{f_2(\mathbf{c}+\mathbf{h}) - A_2(\mathbf{c}+\mathbf{h})}{\|\mathbf{h}\|}, \ldots, \frac{f_n(\mathbf{c}+\mathbf{h}) - A_n(\mathbf{c}+\mathbf{h})}{\|\mathbf{h}\|}\right),$$

and so

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\|R(\mathbf{h}\|}{\|\mathbf{h}\|} = 0, \tag{4.2.5}$$

that is, A is the best affine approximation to f at $\mathbf{c}$, if and only if

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{f_k(\mathbf{c}+\mathbf{h}) - A_k(\mathbf{c}+\mathbf{h})}{\|\mathbf{h}\|} = 0 \tag{4.2.6}$$

for $k = 1, 2, \ldots, n$. But (4.2.6) is the statement that A_k is the best affine approximation to f_k at $\mathbf{c}$. In other words, A is the best affine approximation to f at $\mathbf{c}$ if and only if A_k is the best affine approximation to f_k at $\mathbf{c}$ for $k = 1, 2, \ldots, n$. This result has many interesting consequences.

Proposition If $f_k : \mathbb{R}^m \to \mathbb{R}$ is the kth coordinate function of $f : \mathbb{R}^m \to \mathbb{R}^n$, then f is differentiable at a point $\mathbf{c}$ if and only if f_k is differentiable at $\mathbf{c}$ for $k = 1, 2, \ldots, n$.

Definition If $f_k : \mathbb{R}^m \to \mathbb{R}$ is the kth coordinate function of $f : \mathbb{R}^m \to \mathbb{R}^n$, then we say f is C^1 on an open set U if f_k is C^1 on U for $k = 1, 2, \ldots, n$.

Putting our results in Section 3.3 together with the previous proposition and definition, we have the following basic result.

Theorem If $f : \mathbb{R}^m \to \mathbb{R}^n$ is C^1 on an open ball containing the point $\mathbf{c}$, then f is differentiable at $\mathbf{c}$.

Suppose $f : \mathbb{R}^m \to \mathbb{R}^n$ is differentiable at $\mathbf{c} = (c_1, c_2, \ldots, c_m)$ with best affine approximation A and $f_k : \mathbb{R}^m \to \mathbb{R}$ and $A_k : \mathbb{R}^m \to \mathbb{R}$ are the coordinate functions of f and A, respectively, for $k = 1, 2, \ldots, n$. Since A_k is the best affine approximation to f_k at $\mathbf{c}$, we know from Section 3.3 that

$$A_k(\mathbf{x}) = \nabla f_k(\mathbf{c}) \cdot (\mathbf{x} - \mathbf{c}) + f_k(\mathbf{c}) \tag{4.2.7}$$

for all $\mathbf{x}$ in $\mathbb{R}^m$. Hence, writing the vectors as column vectors, we have

$$\begin{aligned} A(\mathbf{x}) &= \begin{bmatrix} A_1(\mathbf{x}) \\ A_2(\mathbf{x}) \\ \vdots \\ A_n(\mathbf{x}) \end{bmatrix} \\ &= \begin{bmatrix} \nabla f_1(\mathbf{c}) \cdot (\mathbf{x} - \mathbf{c}) + f_1(\mathbf{c}) \\ \nabla f_2(\mathbf{c}) \cdot (\mathbf{x} - \mathbf{c}) + f_2(\mathbf{c}) \\ \vdots \\ \nabla f_n(\mathbf{c}) \cdot (\mathbf{x} - \mathbf{c}) + f_n(\mathbf{c}) \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} \frac{\partial}{\partial x_1} f_1(\mathbf{c}) & \frac{\partial}{\partial x_2} f_1(\mathbf{c}) & \cdots & \frac{\partial}{x_m} f_1(\mathbf{c}) \\ \frac{\partial}{\partial x_1} f_2(\mathbf{c}) & \frac{\partial}{\partial x_2} f_2(\mathbf{c}) & \cdots & \frac{\partial}{x_m} f_2(\mathbf{c}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} f_n(\mathbf{c}) & \frac{\partial}{\partial x_2} f_n(\mathbf{c}) & \cdots & \frac{\partial}{x_m} f_n(\mathbf{c}) \end{bmatrix} \begin{bmatrix} x_1 - c_1 \\ x_2 - c_2 \\ \vdots \\ x_m - c_m \end{bmatrix} + \begin{bmatrix} f_1(\mathbf{c}) \\ f_2(\mathbf{c}) \\ \vdots \\ f_m(\mathbf{c}) \end{bmatrix}. \quad (4.2.8)$$

It follows that the $n \times m$ matrix in (4.2.8) is the derivative of f.

Theorem If $f : \mathbb{R}^m \to \mathbb{R}^n$ is differentiable at a point $\mathbf{c}$, then the derivative of f at $\mathbf{c}$ is given by

$$Df(\mathbf{c}) = \begin{bmatrix} \frac{\partial}{\partial x_1} f_1(\mathbf{c}) & \frac{\partial}{\partial x_2} f_1(\mathbf{c}) & \cdots & \frac{\partial}{x_m} f_1(\mathbf{c}) \\ \frac{\partial}{\partial x_1} f_2(\mathbf{c}) & \frac{\partial}{\partial x_2} f_2(\mathbf{c}) & \cdots & \frac{\partial}{x_m} f_2(\mathbf{c}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} f_n(\mathbf{c}) & \frac{\partial}{\partial x_2} f_n(\mathbf{c}) & \cdots & \frac{\partial}{x_m} f_n(\mathbf{c}) \end{bmatrix}. \quad (4.2.9)$$

We call the matrix in (4.2.9) the *Jacobian matrix* of f, after the German mathematician Carl Gustav Jacob Jacobi (1804-1851). Note that we have seen this matrix before in our discussion of change of variables in integrals in Section 3.7.

Example Consider the function $f : \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$f(x, y, z) = (xyz, 3x - 2yz).$$

The coordinate functions of f are

$$f_1(x, y, z) = xyz$$

and

$$f_2(x, y, z) = 3x - 2yz.$$

Now

$$\nabla f_1(x, y, z) = (yz, xz, xy)$$

and

$$\nabla f_2(x, y, z) = (3, -2z, -2y),$$

so the Jacobian of f is

$$Df(x, y, z) = \begin{bmatrix} yz & xz & xy \\ 3 & -2z & -2y \end{bmatrix}.$$

Hence, for example,

$$Df(1, 2, -1) = \begin{bmatrix} -2 & -1 & 2 \\ 3 & 2 & -4 \end{bmatrix}.$$

Since $f(1,2,-1) = (-2,7)$, the best affine approximation to f at $(1,2,-1)$ is

$$\begin{aligned} A(x,y,z) &= \begin{bmatrix} -2 & -1 & 2 \\ 3 & 2 & -4 \end{bmatrix} \begin{bmatrix} x-1 \\ y-2 \\ z+1 \end{bmatrix} + \begin{bmatrix} -2 \\ 7 \end{bmatrix} \\ &= \begin{bmatrix} -2(x-1)-(y-2)+2(z+1)-2 \\ 3(x-1)+2(y-2)-4(z+1)+7 \end{bmatrix} \\ &= \begin{bmatrix} -2x-y+2z+4 \\ 3x+2y-4z-4 \end{bmatrix}. \end{aligned}$$

Tangent planes

Suppose $f : \mathbb{R}^2 \to \mathbb{R}^3$ parametrizes a surface S in $\mathbb{R}^3$. If f_1, f_2, and f_3 are the coordinate functions of f, then the best affine approximation to f at a point (s_0, t_0) is given by

$$\begin{aligned} A(s,t) &= \begin{bmatrix} \frac{\partial}{\partial s} f_1(t_0,s_0) & \frac{\partial}{\partial t} f_1(t_0,s_0) \\ \frac{\partial}{\partial s} f_2(t_0,s_0) & \frac{\partial}{\partial t} f_2(t_0,s_0) \\ \frac{\partial}{\partial s} f_3(t_0,s_0) & \frac{\partial}{\partial t} f_3(t_0,s_0) \end{bmatrix} \begin{bmatrix} s-s_0 \\ t-t_0 \end{bmatrix} + \begin{bmatrix} f_1(s_0,t_0) \\ f_2(s_0,t_0) \\ f_3(s_0,t_0) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial}{\partial s} f_1(s_0,t_0) \\ \frac{\partial}{\partial s} f_2(s_0,t_0) \\ \frac{\partial}{\partial s} f_3(s_0,t_0) \end{bmatrix} (s-s_0) + \begin{bmatrix} \frac{\partial}{\partial t} f_1(s_0,t_0) \\ \frac{\partial}{\partial t} f_2(s_0,t_0) \\ \frac{\partial}{\partial t} f_3(s_0,t_0) \end{bmatrix} (t-t_0) + \begin{bmatrix} f_1(s_0,t_0) \\ f_2(s_0,t_0) \\ f_3(s_0,t_0) \end{bmatrix} \end{aligned} \tag{4.2.10}$$

If the vectors

$$\mathbf{v} = \begin{bmatrix} \frac{\partial}{\partial s} f_1(s_0,t_0) \\ \frac{\partial}{\partial s} f_2(s_0,t_0) \\ \frac{\partial}{\partial s} f_3(s_0,t_0) \end{bmatrix} \tag{4.2.11}$$

and

$$\mathbf{w} = \begin{bmatrix} \frac{\partial}{\partial t} f_1(s_0,t_0) \\ \frac{\partial}{\partial t} f_2(s_0,t_0) \\ \frac{\partial}{\partial t} f_3(s_0,t_0) \end{bmatrix} \tag{4.2.12}$$

are linearly independent, then (4.2.10) implies that the image of A is a plane in $\mathbb{R}^3$ which passes through the point $f(s_0, t_0)$ on the surface S. Moreover, if we let C_1 be the curve on S through the point $f(s_0, t_0)$ parametrized by $\varphi_1(s) = f(s, t_0)$ and C_2 be the curve on S through the point $f(s_0, t_0)$ parametrized by $\varphi_2(t) = f(s_0, t)$, then $\mathbf{v}$ is tangent to C_1 at $f(s_0, t_0)$ and $\mathbf{w}$ is tangent to C_2 at $f(s_0, t_0)$. Hence we call the image of A the *tangent plane* to the surface S at the point $f(s_0, t_0)$.

Example Let T be the torus parametrized by

$$f(s,t) = ((3+\cos(t))\cos(s), (3+\cos(t))\sin(s), \sin(t))$$

for $0 \leq s \leq 2\pi$ and $0 \leq t \leq 2\pi$. Then

$$Df(s,t) = \begin{bmatrix} -(3+\cos(t))\sin(s) & -\sin(t)\cos(s) \\ (3+\cos(t))\cos(s) & -\sin(t)\sin(s) \\ 0 & \cos(t) \end{bmatrix}.$$

Thus, for example,

$$Df\left(\frac{\pi}{2}, \frac{\pi}{4}\right) = \begin{bmatrix} -\left(3+\dfrac{1}{\sqrt{2}}\right) & 0 \\ 0 & -\dfrac{1}{\sqrt{2}} \\ 0 & \dfrac{1}{\sqrt{2}} \end{bmatrix}.$$

Since

$$f\left(\frac{\pi}{2}, \frac{\pi}{4}\right) = \left(0, 3+\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right),$$

the best affine approximation to f at $\left(\frac{\pi}{2}, \frac{\pi}{4}\right)$ is

$$\begin{aligned} A(s,t) &= \begin{bmatrix} -\left(3+\dfrac{1}{\sqrt{2}}\right) & 0 \\ 0 & -\dfrac{1}{\sqrt{2}} \\ 0 & \dfrac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} s-\dfrac{\pi}{2} \\ t-\dfrac{\pi}{4} \end{bmatrix} + \begin{bmatrix} 0 \\ 3+\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \end{bmatrix} \\ &= \begin{bmatrix} -\left(3+\dfrac{1}{\sqrt{2}}\right) \\ 0 \\ 0 \end{bmatrix} \left(s-\frac{\pi}{2}\right) + \begin{bmatrix} 0 \\ -\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \end{bmatrix} \left(t-\frac{\pi}{4}\right) + \begin{bmatrix} 0 \\ 3+\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \end{bmatrix}. \end{aligned}$$

Hence

$$\begin{aligned} x &= -\left(3+\frac{1}{\sqrt{2}}\right)\left(s-\frac{\pi}{2}\right), \\ y &= -\frac{1}{\sqrt{2}}\left(t-\frac{\pi}{4}\right) + 3 + \frac{1}{\sqrt{2}}, \\ z &= \frac{1}{\sqrt{2}}\left(t-\frac{\pi}{4}\right) + \frac{1}{\sqrt{2}}, \end{aligned}$$

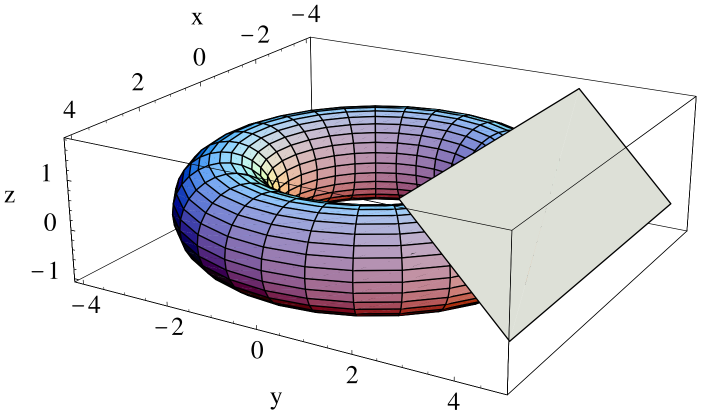

Figure 4.2.1 Torus with a tangent plane

are parametric equations for the plane P tangent to T at $\left(0, 3+\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$. See Figure 4.2.1.

Chain rule

We are now in a position to state the chain rule in its most general form. Consider functions $g : \mathbb{R}^m \to \mathbb{R}^q$ and $f : \mathbb{R}^q \to \mathbb{R}^n$ and suppose g is differentiable at $\mathbf{c}$ and f is differentiable at $g(\mathbf{c})$. Let $h : \mathbb{R}^m \to \mathbb{R}^n$ be the composition $h(\mathbf{x}) = f(g(\mathbf{x}))$ and denote the coordinate functions of f, g, and h by f_i, $i = 1, 2, \ldots, n$, g_j, $j = 1, 2 \ldots, q$, and h_k, $k = 1, 2, \ldots, n$, respectively. Then, for $k = 1, 2, \ldots, n$,

$$h_k(x_1, x_2, \ldots, x_m) = f_k(g_1(x_1, x_2, \ldots, x_m), g_2(x_1, x_2, \ldots, x_m), \ldots, g_q(x_1, x_2, \ldots, x_m)).$$

Now if we fix $m - 1$ of the variables x_1, x_2, $\ldots$, x_m, say, all but x_j, then h_k is the composition of a function from $\mathbb{R}$ to $\mathbb{R}^q$ with a function from $\mathbb{R}^q$ to $\mathbb{R}$. Thus we may use the chain rule from Section 3.3 to compute $\frac{\partial}{\partial x_j} h_k(\mathbf{c})$, namely,

$$\begin{aligned} \frac{\partial}{\partial x_j} h_k(\mathbf{c}) &= \nabla f_k(g(\mathbf{c})) \cdot \left(\frac{\partial}{\partial x_j} g_1(\mathbf{c}), \frac{\partial}{\partial x_j} g_2(\mathbf{c}), \ldots, \frac{\partial}{x_j} g_q(\mathbf{c}) \right) \\ &= \frac{\partial}{\partial x_1} f_k(g(\mathbf{c})) \frac{\partial}{\partial x_j} g_1(\mathbf{c}) + \frac{\partial}{\partial x_2} f_k(g(\mathbf{c})) \frac{\partial}{\partial x_j} g_2(\mathbf{c}) + \\ &\qquad \cdots + \frac{\partial}{\partial x_q} f_k(g(\mathbf{c})) \frac{\partial}{\partial x_j} g_q(\mathbf{c}). \end{aligned} \tag{4.2.13}$$

Hence $\frac{\partial}{\partial x_j} h_k(\mathbf{c})$ is equal to the dot product of the kth row of $Df(g(\mathbf{c}))$ with the jth column of $Dg(\mathbf{c})$. Moreover, if g is C^1 on an open ball about $\mathbf{c}$ and f is C^1 on an open ball about $g(\mathbf{c})$, then (4.2.13) shows that $\frac{\partial}{\partial x_j} h_k$ is continuous on an open ball about $\mathbf{c}$. It follows from our results in Section 3.3 that h is differentiable at $\mathbf{c}$. Since $\frac{\partial}{\partial x_j} h_k$ is the entry in the kth row and jth column of $Dh(\mathbf{c})$, (4.2.13) implies $Dh(\mathbf{c}) = Df(g(\mathbf{c}))Dg(\mathbf{c})$. This result, the chain rule, may be proven without assuming that f and g are both C^1, and so we state the more general result in the following theorem.

Chain Rule If $g : \mathbb{R}^m \to \mathbb{R}^q$ is differentiable at $\mathbf{c}$ and $f : \mathbb{R}^q \to \mathbb{R}^n$ is differentiable at $g(\mathbf{c})$, then $f \circ g$ is differentiable at $\mathbf{c}$ and

$$D(f \circ g)(\mathbf{c}) = Df(g(\mathbf{c}))Dg(\mathbf{c}). \tag{4.2.14}$$

Equivalently, the chain rule says that if A is the best affine approximation to g at $\mathbf{c}$ and B is the best affine approximation to f at $g(\mathbf{c})$, then $B \circ A$ is the best affine approximation to $f \circ g$ at $\mathbf{c}$. That is, the best affine approximation to a composition of functions is the composition of the individual best affine approximations.

Example Suppose $g : \mathbb{R}^2 \to \mathbb{R}^3$ is defined by

$$g(s,t) = (\cos(s)\sin(t), \sin(s)\sin(t), \cos(t))$$

and $f : \mathbb{R}^3 \to \mathbb{R}^2$ is defined by

$$f(x,y,z) = (10xyz, x^2 - yz).$$

Then

$$Dg(s,t) = \begin{bmatrix} -\sin(s)\sin(t) & \cos(s)\cos(t) \\ \cos(s)\sin(t) & \sin(s)\cos(t) \\ 0 & -\sin(t) \end{bmatrix}$$

and

$$Df(x,y,z) = \begin{bmatrix} 10yz & 10xz & 10xy \\ 2x & -z & -y \end{bmatrix}.$$

Let $h(s,t) = f(g(s,t))$. To find $Dh\left(\frac{\pi}{4}, \frac{\pi}{4}\right)$, we first note that

$$g\left(\frac{\pi}{4}, \frac{\pi}{4}\right) = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}\right),$$

$$Dg\left(\frac{\pi}{4}, \frac{\pi}{4}\right) = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \\ 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

and

$$Df\left(g\left(\frac{\pi}{4}, \frac{\pi}{4}\right)\right) = Df\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}\right) = \begin{bmatrix} \frac{5}{\sqrt{2}} & \frac{5}{\sqrt{2}} & \frac{5}{2} \\ 1 & -\frac{1}{\sqrt{2}} & -\frac{1}{2} \end{bmatrix}.$$

Thus

$$\begin{aligned} Dh\left(\frac{\pi}{4},\frac{\pi}{4}\right) &= Df\left(g\left(\frac{\pi}{4},\frac{\pi}{4}\right)\right) Dg\left(\frac{\pi}{4},\frac{\pi}{4}\right)\\ &= \begin{bmatrix} \frac{5}{\sqrt{2}} & \frac{5}{\sqrt{2}} & \frac{5}{2} \\ 1 & -\frac{1}{\sqrt{2}} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \\ 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}\\ &= \begin{bmatrix} 0 & \frac{5}{2\sqrt{2}} \\ -\frac{1+\sqrt{2}}{2\sqrt{2}} & \frac{1}{2} \end{bmatrix}. \end{aligned}$$

Problems

1. Find the best affine approximation for each of the following functions at the specified point $\mathbf{c}$.
   (a) $f(x,y) = (x^2+y^2, 3xy)$, $\mathbf{c} = (1,2)$
   (b) $g(x,y,z) = (\sin(x+y+z), xy\cos(z))$, $\mathbf{c} = \left(0, \frac{\pi}{4}, \frac{\pi}{4}\right)$
   (c) $h(s,t) = (3s^2+t, s-t, 4st^2, 4t-s)$, $\mathbf{c} = (-1,3)$
2. Each of the following functions parametrizes a surface S in $\mathbb{R}^3$. In each case, find parametric equations for the tangent plane P passing through the point $f(s_0,t_0)$. Plot S and P together.
   (a) $f(s,t) = (t\cos(s), t\sin(s), t)$, $(s_0,t_0) = \left(\frac{\pi}{2}, 2\right)$
   (b) $f(s,t) = (t^2\cos(s), t^2, t^2\sin(s))$, $(s_0,t_0) = (0,1)$
   (c) $f(s,t) = (\cos(s)\sin(t), \sin(s)\sin(t), \cos(t))$, $(s_0,t_0) = \left(\frac{\pi}{2}, \frac{\pi}{4}\right)$
   (d) $f(s,t) = (3\cos(s)\sin(t), \sin(s)\sin(t), 2\cos(t))$, $(s_0,t_0) = \left(\frac{\pi}{4}, \frac{\pi}{4}\right)$
   (e) $f(s,t) = ((4+2\cos(t))\cos(s), (4+2\cos(t))\sin(s), 2\sin(t))$, $(s_0,t_0) = \left(\frac{3\pi}{4}, \frac{\pi}{4}\right)$
3. Let S be the graph of a function $f : \mathbb{R}^2 \to \mathbb{R}$. Define the function $F : \mathbb{R}^2 \to \mathbb{R}^3$ by $F(s,t) = (s,t,f(s,t))$. We may find an equation for the plane tangent to S at $(s_0,t_0,f(s_0,t_0))$ using either the techniques of Section 3.3 (looking at S as the graph of f) or the techniques of this section (looking at S as a surface parametrized by F). Verify that these two approaches yield equations for the same plane, both in the special case when $f(s,t) = s^2+t^2$ and $(s_0,t_0) = (1,2)$, and in the general case.
4. Use the chain rule to find the derivative of $f \circ g$ at the point $\mathbf{c}$ for each of the following.
   (a) $f(x,y) = (x^2y, x-y)$, $g(s,t) = (3st, s^2-4t)$, $\mathbf{c} = (1,-2)$
   (b) $f(x,y,z) = (4xy, 3xz)$, $g(s,t) = \left(st^2-4t, s^2, \dfrac{4}{st}\right)$, $\mathbf{c} = (-2,3)$
   (c) $f(x,y) = (3x+4y, 2x^2y, x-y)$, $g(s,t,u) = (4s-3t+u, 5st^2)$, $\mathbf{c} = (1,-2,3)$

5. Suppose

$$x = f(u,v),$$
$$y = g(u,v),$$

and

$$u = h(s,t),$$
$$v = k(s,t).$$

(a) Show that

$$\frac{\partial x}{\partial s} = \frac{\partial x}{\partial u}\frac{\partial u}{\partial s} + \frac{\partial x}{\partial v}\frac{\partial v}{\partial s}$$

and

$$\frac{\partial x}{\partial t} = \frac{\partial x}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial x}{\partial v}\frac{\partial v}{\partial t}.$$

(b) Find similar expressions for $\frac{\partial y}{\partial s}$ and $\frac{\partial y}{\partial t}$.

6. Use your results in Problem 5 to find $\frac{\partial x}{\partial s}$, $\frac{\partial x}{\partial t}$, $\frac{\partial y}{\partial s}$, and $\frac{\partial y}{\partial t}$ when

$$x = u^2 v,$$
$$y = 3u - v,$$

and

$$u = 4t^2 - s^2,$$
$$v = \frac{4t}{s}.$$

7. Suppose T is a function of x and y where

$$x = r\cos(\theta),$$
$$y = r\sin(\theta).$$

Show that

$$\frac{\partial T}{\partial r} = \frac{\partial T}{\partial x}\cos(\theta) + \frac{\partial T}{\partial y}\sin(\theta)$$

and

$$\frac{\partial T}{\partial \theta} = -\frac{\partial T}{\partial x}r\sin(\theta) + \frac{\partial T}{\partial y}r\cos(\theta).$$

8. Suppose the temperature at a point (x,y) in the plane is given by

$$T = 100 - \frac{20}{\sqrt{1+x^2+y^2}}.$$

(a) If (r,θ) represents the polar coordinates of (x,y), use Problem 7 to find $\frac{\partial T}{\partial r}$ and $\frac{\partial T}{\partial \theta}$ when $r = 4$ and $\theta = \frac{\pi}{6}$.

(b) Show that $\frac{\partial T}{\partial \theta} = 0$ for all values of r and θ. Can you explain this result geometrically?

9. Let T be the torus parametrized by

$$\begin{aligned} x &= (4 + 2\cos(t))\cos(s), \\ y &= (4 + 2\cos(t))\sin(s), \\ z &= 2\sin(t), \end{aligned}$$

for $0 \leq s \leq 2\pi$ and $0 \leq t \leq 2\pi$.

(a) If U is a function of x, y, and z, find general expressions for $\frac{\partial U}{\partial s}$ and $\frac{\partial U}{\partial t}$.

(b) Suppose

$$U = 80 - 40e^{-\frac{1}{20}(x^2+y^2+z^2)}$$

gives the temperature at a point (x, y, z) on T. Find expressions for $\frac{\partial U}{\partial s}$ and $\frac{\partial U}{\partial t}$ in this case. What is the geometrical interpretation of these quantities?

(c) Evaluate $\frac{\partial U}{\partial s}$ and $\frac{\partial U}{\partial t}$ in the particular case $s = \frac{\pi}{4}$ and $t = \frac{\pi}{4}$.