Chapter 3

Differential Calculus

3.1 The Jacobi matrix and linearization

Key Ideas.

- *Recall Taylor's theorem*

$$f(t_* + \Delta t) = f(t_*) + f'(t_*)\Delta t + \textit{remainder}$$

  The best linear approximation of f *near* t_* *is*

$$L(\Delta t) = f(t_*) + f'(t_*)\Delta t.$$

- *Multivariable version*

$$\begin{aligned} &T(u_* + \Delta u, v_* + \Delta v, \dots) \\ &\quad = T(u_*, v_*, \dots) + \begin{pmatrix} | & | & \\ \frac{\partial T}{\partial u}(u_*, v_*, \dots) & \frac{\partial T}{\partial v}(u_*, v_*, \dots) & \cdots \\ | & | & \end{pmatrix} \begin{pmatrix} \Delta u \\ \Delta v \\ \vdots \end{pmatrix} \\ &\qquad\qquad + \textit{remainder} \end{aligned}$$

  Write as

$$T(u_*+\Delta u, v_*+\Delta v, \dots) = T(u_*, v_*, \dots) + DT(u_*, v_*, \dots) \begin{pmatrix} \Delta u \\ \Delta v \\ \vdots \end{pmatrix} + \textit{remainder}$$

- *The multivariable derivative is the* ***Jacobi matrix***

$$DT = \begin{pmatrix} | & | & \\ \frac{\partial T}{\partial u} & \frac{\partial T}{\partial v} & \cdots \\ | & | & \end{pmatrix} = \begin{pmatrix} | & | & \\ \partial_u & \partial_v & \cdots \\ | & | & \end{pmatrix}$$

  If we write $T(u, v, \dots) = (x(u, v, \dots), y(u, v, \dots), \dots)$ *then*

$$DT = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \cdots \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \cdots \\ \vdots & \vdots & \end{pmatrix}$$

- *The best linear approximation of* T *at the point* (u_*, v_*) *is given by the affine transformation*

$$L(\Delta u, \Delta v) = T(u_*, v_*) + DT(u_*, v_*) \begin{pmatrix} \Delta u \\ \Delta v \end{pmatrix}$$

Exercises.

1. Consider the transformation

$$T(u,v) = (u\cos(v), u\sin(v)).$$

   (a) Draw, as best you can, the transformation T.
   (b) Compute the Jacobi matrix for T.
   (c) Use the Jacobi matrix to find the best linear approximation of T at the point $(u,v) = (1, \pi/3)$.
   (d) Draw, as best you can, the linear approximation. Be sure to indicate the coordinate vector fields in your picture.

2. Consider the transformation

$$T(u,v) = (u\cos(v), u\sin(v), u^2).$$

   (a) Draw, as best you can, the transformation T.
   (b) Compute the Jacobi matrix for T.
   (c) Use the Jacobi matrix to find the best linear approximation of T at the point $(u,v) = (1, \pi/4)$.
   (d) Draw, as best you can, the linear approximation. Be sure to indicate the coordinate vector fields in your picture.

3. Consider the transformation

$$T(u,v) = (\cos(v), \sin(v), u).$$

   (a) Draw, as best you can, the transformation T.
   (b) Compute the Jacobi matrix for T.
   (c) Use the Jacobi matrix to find the best linear approximation of T at the point $(u,v) = (1, \pi/4)$.
   (d) Draw, as best you can, the linear approximation. Be sure to indicate the coordinate vector fields in your picture.

4. Consider the transformation

$$T(t) = (\cos(t), \sin(t)).$$

   (a) Draw, as best you can, the transformation T.
   (b) Compute the Jacobi matrix for T.
   (c) Use the Jacobi matrix to find the best linear approximation of T at the point $t = \pi/4$.
   (d) Draw, as best you can, the linear approximation. Be sure to indicate the coordinate vector fields in your picture.

5. Consider the transformation

$$T(t) = (\cos(t), \sin(t)).$$

   (a) Draw, as best you can, the transformation T.
   (b) Compute the Jacobi matrix for T.
   (c) Use the Jacobi matrix to find the best linear approximation of T at the point $t = \pi/4$.
   (d) Draw, as best you can, the linear approximation. Be sure to indicate the coordinate vector fields in your picture.

6. Consider the transformation

$$T(x, y) = x^2 - y^2.$$

   (a) Draw, as best you can, the transformation T.
   (b) Compute the Jacobi matrix for T.
   (c) Use the Jacobi matrix to find the best linear approximation of T at the point $(x, y) = (2, 3)$.
   (d) Draw, as best you can, the linear approximation. Be sure to indicate the coordinate vector fields in your picture.

7. Consider the transformation

$$T(u, v) = (u^2 - v^2, 2uv),$$

   where $u \geq 0$.

   (a) Draw, as best you can, the transformation T.
   (b) Compute the Jacobi matrix for T.
   (c) Use the Jacobi matrix to find the best linear approximation of T at the point $(u, v) = (1, 1)$.
   (d) Draw, as best you can, the linear approximation. Be sure to indicate the coordinate vector fields in your picture.

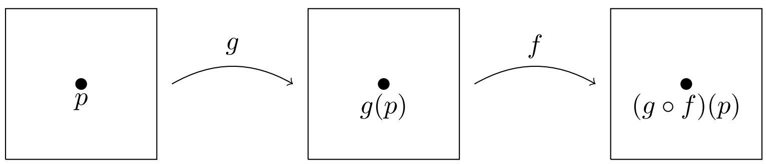
3.2 The Chain Rule

Key Ideas.

- *Chain rule from Calculus 1:*

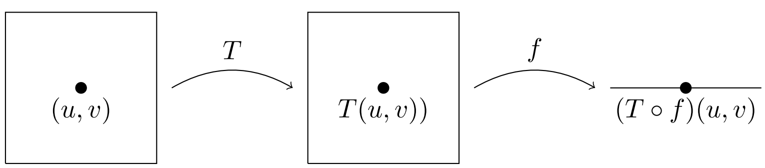
$$\frac{d}{dt}(f \circ g)(t) = f'(g(t))\, g'(t).$$

- *General situation*



$$D(f \circ g)(p) = Df(g(p))\, Dg(p)$$

- *Special case: coordinate changes and scalar functions*



Suppose $(x, y) = T(u, v)$*. Then* $D(f \circ T)(u, v) = Df(x, y)DT(u, v)$ *becomes*

$$\begin{pmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}$$

We write this as

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u}, \qquad \textit{etc.}$$

- *Important example: polar coordinates*

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

Exercises.

1. Let T_1 and T_2 be two transformations of the Euclidean plane with:

$$T_1(0,1) = (1,1), \quad T_1(1,1) = (0,1),$$
$$T_2(0,1) = (0,1), \quad T_2(1,1) = (1,1).$$

It is furthermore known that the Jacobi matrices of T_1 and T_2 at points $(0,1)$ and $(1,1)$ are:

$$DT_1\big|_{(0,1)} = \begin{pmatrix} -1 & 2 \\ -2 & 1 \end{pmatrix}, \quad DT_1\big|_{(1,1)} = \begin{pmatrix} 0 & 3 \\ 5 & 1 \end{pmatrix},$$
$$DT_2\big|_{(0,1)} = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}, \quad DT_2\big|_{(1,1)} = \begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix}.$$

Find the following Jacobi matrices:

(a) $D(T_2 \circ T_1)\big|_{(0,1)}$

(b) $D(T_2 \circ T_1)\big|_{(1,1)}$

(c) $D(T_1 \circ T_2)\big|_{(0,1)}$

(d) $D(T_1 \circ T_2)\big|_{(1,1)}$

2. Use the Chain Rule to compute $\frac{\partial}{\partial u}(f \circ T)$ and $\frac{\partial}{\partial v}(f \circ T)$ for the following choices of f and T. Express your final answer using u and v variables only.

(a) $f(x,y) = x^2 + y^2$ and

$$(x,y) = T(u,v) = (2u+v, u+2v).$$

(b) $f(x,y,z) = xyz$ and

$$(x,y,z) = T(u,v,w) = (u+v+w, v+w, w).$$

(c) $f(x,y,z) = x^2 + y^2 + z^2$ and

$$(x,y,z) = T(u,v) = (u^2, u\sin(v), u\cos(v)).$$

3. Consider the transformation

$$(x,y) = T(u,v) = (u+v, -u+v),$$

and let $f(x,y)$ be some function.

(a) Express $\frac{\partial f}{\partial u}$ and $\frac{\partial f}{\partial v}$ in terms of $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$.

(b) Express $\frac{\partial^2 f}{\partial u^2}$, $\frac{\partial^2 f}{\partial u \partial v}$ and $\frac{\partial^2 f}{\partial v^2}$ in terms of partial derivatives with respect to (x,y)-variables.
[Hint: $\frac{\partial^2 f}{\partial u^2} = \frac{\partial}{\partial u}\frac{\partial f}{\partial u}$, etc.]

4. Let (r, θ) denote the standard polar coordinates in the xy-plane.
   (a) Express $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ in terms of $\frac{\partial f}{\partial r}$ and $\frac{\partial f}{\partial \theta}$.
   (b) Express $\frac{\partial^2 f}{\partial x^2}$, $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y^2}$ in terms of $\frac{\partial^2 f}{\partial r^2}$, $\frac{\partial^2 f}{\partial r \partial \theta}$ and $\frac{\partial^2 f}{\partial \theta^2}$.
5. Let (r, θ, ϕ) denote the standard spherical coordinates in the xyz-space. Let $f(x, y, z)$ be some function which depends only on r:

$$f(x, y, z) = f(r), \quad \frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial \phi} = 0.$$

   (a) Express $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial z}$ in terms of $\frac{\partial f}{\partial r}$.
   (b) Express $\frac{\partial^2 f}{\partial x^2}$, $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y^2}$ in terms of $\frac{\partial^2 f}{\partial r^2}$, $\frac{\partial f}{\partial r}$ etc.

3.3 The second derivative, the Hessian and the second order Taylor approximation

Key Ideas.

- *Suppose* $f\colon \mathbb{R}^2 \to \mathbb{R}$ *and* (x_*, y_*) *is a fixed point.*
- *Then* $Df\colon \mathbb{R}^2 \to \mathbb{R}^2$ *is given by*

$$Df(x,y) = \begin{pmatrix} \frac{\partial f}{\partial x}(x,y) & \frac{\partial f}{\partial y}(x,y) \end{pmatrix}.$$

- *The second derivative* $D(Df) = D^2 f$ *is given by*

$$D^2 f(x,y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial x \partial y}(x,y) \\ \frac{\partial^2 f}{\partial x \partial y}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y) \end{pmatrix}.$$

  This matrix is called the ***Hessian of f****.*

- *We organize the quadratic approximation (second order Taylor expansion) at* (x_*, y_*) *by*

$$\begin{aligned} f(x_* + \Delta x, y_* + \Delta y) & \\ &= f(x_*, y_*) + Df(x_*, y_*) \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \\ &\quad + \begin{pmatrix} \Delta x & \Delta y \end{pmatrix} D^2 f(x_*, y_*) \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} + \textit{remainder} \end{aligned}$$

Exercises.

1. This problem concerns the function $f(x) = \ln(x)$.
   (a) Find the tangent line to the graph of the function f at the point $x_0 = 1$.
   (b) Find the quadratic (second order) Taylor approximation of the function f at $x_0 = 1$.
   (c) Use ‘technology to plot a graph of the function, its tangent line, and the second order Taylor approximation that you found. Focus on the vicinity of the domain point $x_0 = 1$. Use technology to generate a plot showing that the second order Taylor approximations does a better job of approximating the function f.
2. Repeat the previous problem for the function $f(x) = \sqrt{1+4x}$ near $x_0 = 0$.
3. Compute the Jacobian and the Hessian of the following functions.
   (a) $f(x,y) = 2x + 3y$;

(b) $f(x,y) = e^{-x^2-y^2}$;

(c) $f(x,y) = \arctan\left(\frac{y}{x}\right)$;

(d) $f(x,y,z) = \frac{1}{\sqrt{x^2+y^2+z^2}}$.

4. This problem concerns the function

$$f(x,y) = e^{-\frac{x^2+y^2}{2}}.$$

   (a) Find the tangent plane to the graph of the function f at the point $(0,0)$.

   (b) Find the quadratic (Taylor) approximation of the function f at $(0,0)$.

   (c) Use "technology" to graph the function, its tangent plane (linearization), and its quadratic (Taylor) approximation in the vicinity of the domain point $(0,0)$. Provide a print-out which clearly demonstrates that the quadratic approximation improves the linear approximation of f.

   (d) Based on the quadratic approximation you found estimate the value of $f(0.1, 0.2)$.

   (e) Repeat the previous question for $f(-0.02, 0.01)$.

5. This problem concerns the function

$$f_2(x,y) = xy - \frac{x^4}{16} - \frac{y^4}{16}.$$

   (a) Find the tangent plane to the graph of the function f_2 at the point $(2,2)$.

   (b) Find the quadratic (Taylor) approximation of the function f_2 at $(2,2)$.

   (c) Use "technology" to graph the function, its tangent plane (linearization), and its quadratic (Taylor) approximation in the vicinity of the domain point $(2,2)$. Provide a print-out which clearly demonstrates that the quadratic approximation improves the linear approximation of f_2.

   (d) Based on the quadratic approximation you found estimate the value of $f_2(2.2, 1.8)$.

   (e) Repeat the previous question for $f_2(1.98, 2.04)$.

6. Find the following quadratic approximations:

   (a) Quadratic approximation of the function $f(x,y) = \ln(x^2+y^2)$ near the point $(1,0)$;

   (b) Quadratic approximation of the function $f(x,y,z) = x^2y^2z^2$ near the point $(1,-1,1)$.

3.4 Gradients and optimization

Key Ideas.

- *The **gradient vector field** $\mathbf{grad}\, f$ is defined by*

$$\operatorname{grad} f = (Df)^t = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \dots \right\rangle = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \vdots \end{pmatrix}$$

- *The gradient is orthogonal to the level curves of f – see the exercises!*
- *Notation: In physics land you often see the symbol $\vec{\nabla} f$ used for $\operatorname{grad} f$.*
- *Suppose $\hat{V}$ is a unit vector. The **directional derivative** of f in direction $\hat{V}$ is given by $D_{\hat{V}} f = \operatorname{grad} f \cdot \hat{V}$.*
- *Application to optimization:*
  - *Input (x_*, y_*) is a **critical point** if $\operatorname{grad} f(x_*, y_*) = 0$.*
  - *The Hessian tells us whether a critical point (x_*, y_*) is a local minimum, a local maximum, or a saddle.*
- *In order to find optimizers along a boundary or constrained set, we use **constrained optimization**:*
  - *The goal is to optimize the function f along the set $C = 0$.*
  - *This happens when $\operatorname{grad} f$ is parallel to $\operatorname{grad} C$, meaning that for some constant λ we have $\operatorname{grad} f = \lambda \operatorname{grad} C$*
  - *Combining this with $C = 0$, we get an algebraic system to solve.*

Exercises.

1. In this exercise you show that the gradient vector field $\operatorname{grad} f$ is always orthogonal to the level curves of f.

   (a) Before we begin consider an example. Let $g(x, y) = x^2 - y^2$. Make a plot that shows both the level curves of g and the gradient vector field. Does it look like the gradient is orthogonal to the level curves? (It might help to make sure that the aspect ratio of your plot is equal to one.)

   (b) Now we consider a generic setup. We suppose that we have a function $f\colon \mathbb{R}^2 \to \mathbb{R}$. We suppose also that we have a path $(x(t), y(t))$ that follows a level curve of f. Draw a diagram that shows this setup.

   (c) Explain why we know that $f(x(t), y(t))$ is a constant. What does this tell us about $\frac{d}{dt}\left(f(x(t), y(t))\right)$?

(d) Now we use the chain rule to compute $\frac{d}{dt}\left(f(x(t),y(t))\right)$. Show that the the chain rule gives us that

$$\frac{d}{dt}\left(f(x(t),y(t))\right) = \text{grad}\, f(x(t),y(t)) \cdot \langle x'(t), y'(t)\rangle.$$

(e) How do the previous two parts combine to tell us that grad f is orthogonal to the level curves of f?

(f) Finally, since grad f is orthogonal to the level curves, it either points "directly uphill" (meaning in the direction where f increases the most) or "directly downhill" (meaning in the direction where f decreases the most). Which is it? What does it mean to "follow the gradient"?

2. Consider the function $f(x,y) = x^2 - 4y^2$.

   (a) Make a plot showing the contours and the gradient vector field for the function f.

   (b) Compute grad $f(1,1)$. Indicate this on your plot.

   (c) Let $\hat{V} = \langle 1, 0\rangle$. Compute $D_{\hat{V}} f$. Indicate this on your plot.

3. Consider the function $f(x,y) = 3x + 3y - x^3 - y^3$. Restrict attention to the domain $-2 \le x \le 2$, $-2 \le y \le 2$.

   (a) Make a plot showing the contours and the gradient vector field for the function f.

   (b) Compute grad $f(0,0)$. Indicate this on your plot.

   (c) Let $\hat{V} = \langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\rangle$. Compute $D_{\hat{V}} f$. Indicate this on your plot.

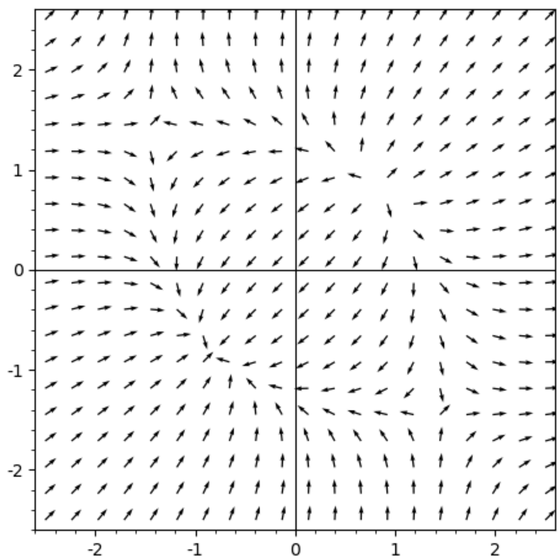
4. Consider the function $f(x,y) = x^2 + 4y^2$.

   (a) Make a plot showing the contours and the gradient vector field for the function f.

   (b) Find a parametrization $(x(t), y(t))$ for the ellipse $x^2 + 4y^2 = 1$.

   (c) Use your parametrization to construct a vector field (parametrized in terms fo t) along the ellipse that everywhere orthogonal to the ellipse.

5. The following plot shows the gradient of some function f. Sketch the contour plot for the function.



Where are the critical points? What type are they?

6. Find all critical points of the following functions and classify them as local maxima, local minima or saddles. Use technology to verify.

   (a) $f(x, y) = 3x^2 - 2x^3 + 3y^2 - 2y^3$;
   (b) $f(x, y) = x^3 - 3xy + y^3$;
   (c) $f(x, y) = x^4 + 4xy - xy^2$;
   (d) $f(x, y) = e^{-x^2-y^2}$.

7. Find the minimum and the maximum of

   (a) The function $f(x, y) = x + y$ subject to the constraint $x^2 + y^2 = 1$;
   (b) The function $f(x, y) = x^2 - y^2$ subject to the constraint $x^2 + y^2 = 1$.

8. Find the point on the ellipse $x^2 + xy + y^2 = 1$ which is

   (a) The closest to the origin;
   (b) The furthest from the origin.

   Based on your findings make a sketch of the ellipse.

9. Find the minimum and the maximum of the function $f(x, y) = x + y$ subject to the constraint $x^2 + y^2 \leq 1$.

10. Consider the function $f(x, y) = x^2 - 2x + y^2$. Find the absolute maximum and the absolute minimum of the function f over the right half-disk

$$x^2 + y^2 \leq 4, \quad x \geq 0.$$