Lecture 11: Chain Rule

Now with partial derivatives, we can learn more tools to study multivariable functions.

1 Total Differential

Total differential describes the rate of change of the value of a function in terms of all variables.
Suppose there is a function $f(x, y, z)$:

$$\begin{aligned} df &= f_x dx + f_y dy + f_z dz \\ &= \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz \end{aligned}$$

Notice that df is not a real number, nor is it Δf. It is an abstract notion. So are dx, dy, and dz.

Question 1. *What is the intuitive way to understand the total differential formula? It is hard to reason about whether it is valid or not.*

Question 2. *If we read the total differential formula from left to right, it seems that we distribute the change in the function value into different dimensions according to partial derivatives. Is this the correct way to understand this formula?*

How do we use the total differential?

- The total differential encodes how changes in independent variables affect the function value.
- We can replace dx, dy, and dz with small variations Δx, Δy, and Δz to get approximated change of the function value, which is the approximation formula:
$$\Delta f \approx f_x \Delta x + f_y \Delta y + f_z \Delta z$$
- We can divide both sides of the total differential by the differential of a common parameter, such as dt, to get the rate of change in terms of this common parameter.

2 Chain Rule

Actually, the third usage of the total differential described above derives the formula of chain rule in multivariable functions.
Suppose there is a function $f(x, y, z)$, and there exists $x = x(t)$, $y = y(t)$, and $z = z(t)$, then

$$\begin{aligned}\frac{df}{dt} &= f_x \frac{dx}{dt} + f_y \frac{dy}{dt} + f_z \frac{dz}{dt} \\ &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}\end{aligned}$$

Why is the formula of chain rule valid?
We can start with the approximation formula. We know that

$$\Delta f \approx f_x \Delta x + f_y \Delta y + f_z \Delta z$$

Then we can divide both sides of the equation by a change of the parameter t, which is Δt:

$$\frac{\Delta f}{\Delta t} \approx f_x \frac{\Delta x}{\Delta t} + f_y \frac{\Delta y}{\Delta t} + f_z \frac{\Delta z}{\Delta t}$$

With $\Delta t \to 0$, according to the definition of derivatives:

$$\begin{aligned}\lim_{\Delta t \to 0} \frac{\Delta f}{\Delta t} &= \frac{df}{dt} \\ \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} &= \frac{dx}{dt} \\ \lim_{\Delta t \to 0} \frac{\Delta y}{\Delta t} &= \frac{dy}{dt} \\ \lim_{\Delta t \to 0} \frac{\Delta z}{\Delta t} &= \frac{dz}{dt}\end{aligned}$$

and the approximation becomes more and more accurate. Eventually

$$\frac{df}{dt} = f_x \frac{dx}{dt} + f_y \frac{dy}{dt} + f_z \frac{dz}{dt}$$

Example 1. *Given that $w = f(x, y, z) = x^2 y + z$, $x = t$, $y = e^t$, and $z = \sin t$, find $\frac{dw}{dt}$.*
Solution:

$$
\begin{aligned}
\frac{dw}{dt} &= f_x \frac{dx}{dt} + f_y \frac{dy}{dt} + f_z \frac{dz}{dt} \\
f_x &= 2xy = 2te^t \\
f_y &= x^2 = t^2 \\
f_z &= 1 \\
\frac{dx}{dt} &= 1 \\
\frac{dy}{dt} &= e^t \\
\frac{dz}{dt} &= \cos t \\
\frac{dw}{dt} &= f_x \frac{dx}{dt} + f_y \frac{dy}{dt} + f_z \frac{dz}{dt} \\
&= 2t \cdot e^t \cdot 1 + t^2 \cdot e^t + 1 \cdot \cos t \\
&= 2te^t + t^2 e^t + \cos t
\end{aligned}
$$

We can verify the result by first find the formula of w with respect to t, and then calculate its derivative:

$$
\begin{aligned}
w &= x^2 y + z \\
&= t^2 e^t + \sin t \\
\frac{dw}{dt} &= \frac{d(t^2 e^t + \sin t)}{dt} \\
&= 2te^t + t^2 e^t + \cos t
\end{aligned}
$$

The results from the two methods are the same.

We can use chain rule of multivariable functions to derive the product and quotient rule of derivatives of single-variable functions:

- Product rule:

$$
\begin{aligned}
f &= u \cdot v \\
\frac{df}{dt} &= \frac{d(u \cdot v)}{dt} \\
&= f_u \frac{du}{dt} + f_v \frac{dv}{dt} \\
&= \frac{du}{dt} v + u \frac{dv}{dt}
\end{aligned}
$$

- Quotient rule:

$$
\begin{aligned}
f &= \frac{u}{v} \\
\frac{df}{dt} &= \frac{d(\frac{u}{v})}{dt} \\
&= f_u \frac{du}{dt} + f_v \frac{dv}{dt} \\
&= \frac{1}{v} \cdot \frac{du}{dt} + \frac{-u}{v^2} \cdot \frac{dv}{dt} \\
&= \frac{\frac{du}{dt} v - u \frac{dv}{dt}}{v^2}
\end{aligned}
$$

3 Chain Rule with More Variables

3.1 Formula of Chain Rule with More Variable

Suppose there is a function $w = f(x, y)$, where $x = x(u, v)$, and $y = y(u, v)$, how to find the total differential of w with respect to u and v?
According to the total differential formula,

$$dw = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

If we apply the total differential formula to the independent variables, then we get

$$
\begin{aligned}
dx &= \frac{\partial x}{\partial u} du + \frac{\partial x}{\partial v} dv \\
dy &= \frac{\partial y}{\partial u} du + \frac{\partial y}{\partial v} dv
\end{aligned}
$$

Therefore

$$
\begin{aligned}
dw &= \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy \\
&= \frac{\partial f}{\partial x} (\frac{\partial x}{\partial u} du + \frac{\partial x}{\partial v} dv) + \frac{\partial f}{\partial y} (\frac{\partial y}{\partial u} du + \frac{\partial y}{\partial v} dv) \\
&= (\frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial u}) du + (\frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial v}) dv
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial w}{\partial u} &= \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial u} \\
\frac{\partial w}{\partial v} &= \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial v}
\end{aligned}
$$

That's the chain rule of multivariable functions where the independent variables are also multivariable functions of some parameters.
Notice that

$$\frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial u} \neq \frac{\partial f}{\partial u}$$

Since they are only part of the change, rather than the full change dx, they cannot be cancelled.

3.2 Intuition behind the Formula

Why the formula is valid?
Suppose that there is a small change in the independent variable u, the change affects the values of both x and y, described by their corresponding partial derivatives. Then the change of x and y affects the value of w eventually, and the relation is described by the total differential formula.
So the chain rule actually describes the propogation relation between independent variables and some function values.

Example 2. *Conversion between Cartesian coordinate system and polar coordinate system.*
Suppose there is a function formula described in Cartesian coordinate system $f(x, y)$.

$$\begin{aligned}
x &= r\cos\theta \\
y &= r\sin\theta \\
\frac{\partial f}{\partial r} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} \\
&= \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta \\
\frac{\partial f}{\partial \theta} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta} \\
&= -r\sin\theta\frac{\partial f}{\partial x} + r\cos\theta\frac{\partial f}{\partial y}
\end{aligned}$$