13.7: The multivariable chain rule

The chain rule with one independent variable

\[ w = f(x, y). \]  
If the particle is moving along a curve \( x = x(t), y = y(t) \), then the values that the particle feels is \( w = f(x(t), y(t)) \). Then, \( w = w(t) \) is a function of \( t \).

\( x, y \) are intermediate variables and \( t \) is the independent variable.

The chain rule says: If both \( f_x \) and \( f_y \) are continuous, then

\[ \frac{dw}{dt} = f_x x'(t) + f_y y'(t). \]

Intuitively, the total change rate is the superposition of contributions in both directions.

**Comment:** You cannot do as in one single variable calculus like

\[ \frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt} = \frac{dw}{dt} + \frac{dw}{dt}, \]

by canceling the two \( dx \)'s. \( \partial w \) in \( \frac{\partial w}{\partial x} \) is only the ‘partial’ change due to the change of \( x \), which is different from the \( dw \) in \( \frac{dw}{dt} \) since the latter is the total change.

The proof is to use the increment \( \Delta w = w_x \Delta x + w_y \Delta y + \text{small} \). Read Page 903.

**Example:** Suppose \( w = \sin(\uv) \) where \( u = 2s \) and \( v = s^2 \). Compute \( \frac{dw}{ds} \) at \( s = 1 \).

**Solution.** By chain rule

\[ \frac{dw}{ds} = \frac{\partial w}{\partial u} \frac{du}{ds} + \frac{\partial w}{\partial v} \frac{dv}{ds} = \cos(uv)v * 2 + \cos(uv)u * 2s. \]

At \( s = 1, u = 2, v = 1 \). Hence, we have

\[ \cos(2) * 2 + \cos(2) * 2 * 2 = 6 \cos(2). \]

We can actually check that \( w = \sin(uv) = \sin(2s^3) \). Hence, \( w'(s) = \cos(2s^3) * 6s^2 \). At \( s = 1 \), this is \( 6 \cos(2) \). \( \square \)

**Exercise:** Derive the formula for \( \frac{d^2 w}{dt^2} \).
Chain rule with several independent variables

\[ w = f(x, y, z) \text{ and } x = x(u, v), y = y(u, v), z = z(u, v). \quad x, y, z \text{ are intermediate variables while } u, v \text{ are independent variables.} \]

\[ w \text{ is then a function of } u, v \text{ through } x, y, z. \text{ The chain rule goes in a similar fashion since when we take partial derivative on } u, \text{ it is just regarded as a constant.} \]

Suppose \( f_x, f_y, f_z \) are continuous, then

\[
\frac{\partial w}{\partial u} = f_x x_u + f_y y_u + f_z z_u.
\]

**Example:** Let \( z = f(u, v) \). \( u = 2x + y, v = 3x - 2y. \) If \( z_u = 3, z_v = -2 \) at \((u, v) = (3, 1)\). Find \( \frac{\partial z}{\partial y} \) at \((x, y) = (1, 1)\). Write \( \frac{\partial^2 z}{\partial y^2} \) in terms of \( f_{uu}, f_{uv} \) and \( f_{vv} \).

**Solution.** At \((x, y) = (1, 1), (u, v) = (3, 1)\). The points given are consistent.

\[
\frac{\partial z}{\partial y} = 3 \frac{\partial u}{\partial y} + (-2) \frac{\partial v}{\partial y} = 3 \cdot 1 + (-2) \cdot (-2) = 7.
\]

For the second, we have found \( \frac{\partial z}{\partial y} = f_u - 2f_v. \) Then,

\[
\frac{\partial^2 z}{\partial y^2} = \frac{\partial f_u}{\partial y} - 2 \frac{\partial f_v}{\partial y} = (f_{uu} - 2f_{uv}) - 2(f_{vu} - 2f_{vv}) = f_{uu} - 4f_{uv} + 4f_{vv}.
\]

Formally, we are just squaring \( \frac{\partial^2 y}{\partial y^2} = (\partial_u - 2\partial_v)^2. \) This would make sense if you know more about operators and the commutativity between them.

When you take partial derivatives by applying chain rules, you really should be clear what variables are being held fixed.

**Example:** consider \( w = x + 2y. \) Let \( z = y - x. \)

- **Regard** \( w \) as a function of \( x \) and \( y. \) Compute \( \frac{\partial w}{\partial x} \) (sometimes we denote it as \( \frac{\partial w}{\partial x} \) or \( \frac{\partial w(x,y)}{\partial x} \))

- **Regard** \( w \) as a function of \( x \) and \( z. \) Compute \( \frac{\partial w}{\partial x} \) (sometimes we denote it as \( \frac{\partial w}{\partial x} \) or \( \frac{\partial w(x,z)}{\partial x} \))
Solution. We only solve the second part as an example. \( w = w(x, y(x, z)) \).

Then,

\[
\frac{\partial w}{\partial x} |_{z} = \frac{\partial w}{\partial x} |_{y} \frac{\partial x}{\partial x} |_{z} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial x} |_{z} = 1 \cdot 1 + 2 \cdot 1 = 3.
\]

This can be verified directly by plugging in \( w = x + 2(z + x) = 3x + 2z \). □

**Example:** Let \( w = xy \ln(u + v) \) where \( u = x + y \) and \( v = (x - y)^{1/2} \).

Compute \( \frac{\partial w}{\partial x} \).

**Implicit differentiation**

**Theorem 1.** Suppose \( F(x_1, x_2, \ldots, x_n, z) \) is continuously differentiable. Consider the level set \( F = C \). If \( F_z(a_1, \ldots, a_n, b) \neq 0 \), then we can solve \( z \) as a function of \( x_1, \ldots, x_n \) near \( (a_1, \ldots, a_n, b) \) such that the function \( z = f(x_1, \ldots, x_n) \) is continuously differentiable and \( f(a_1, \ldots, a_n) = b \).

The proof is omitted. To see this, let us consider the case \( F(x, y) = C \). If \( F_y \neq 0 \), the graph is not vertical. The locally, it is a function graph. The relation \( y = y(x) \) is then an implicit function since the formula is not explicitly given.

What happens if \( F_y = 0 \)? Clearly, the derivative is not defined. \( F_y = 0 \) implies that the tangent line of the level set is parallel with the \( y \) direction. Near the point considered, the curve often is not a function graph. By ‘near’, we mean we can find a small disk that encloses the point we consider.

**Differentiation of implicit functions**

We then have

\[ F(x, y(x)) = C. \]

If we take \( x \) derivative on both sides, we have:

\[ F_x + F_y y'(x) = 0 \Rightarrow y'(x) = -\frac{F_x}{F_y}. \]

**Example:** Consider \( f(x, y) = e^y + y^2 - x - \sin(x) \). Consider the level set passing though \( (0, 0) \). Is it a function graph for \( y = g(x) \) near \( (0, 0) \)? If yes, what is the derivative \( g'(0) \)?

**Solution.**

\[ f_y(0, 0) = e^0 + 2 \cdot 0 = 1 \neq 0. \]
Yes. Then,
\[ g'(0) = -\frac{f_x(0,0)}{f_y(0,0)} = -\frac{-1}{1} = 2. \]

If \( F_x(x_0,y_0) \neq 0 \), we can similarly solve \( x \) in terms of \( y \). For the function \( x = x(y) \),
\[ \frac{dx}{dy} = -\frac{F_y}{F_x}. \]

**Exercise:** Consider \( f(x,y) = \ln(e^x + e^y) - 2x \) and the level curve passing through \((0,0)\). Can we solve \( x \) in terms of \( y \) locally near \((0,0)\)? If yes, what is \( x'(0) \)?

For \( F(x,y,z) = C \) where \( F \) is continuously differentiable, if \( F_z \neq 0 \), then \( z = z(x,y) \).
\[ \frac{\partial z}{\partial x} = -\frac{F_x}{F_z}, \frac{\partial z}{\partial y} = -\frac{F_y}{F_z} \]

The second order derivatives become more involved. However, in principle, you can just take derivatives on the expressions you have already and regard \( z \) as functions of \( x,y \).

**Exercise:** Derive the formula for \( \frac{\partial^2 z}{\partial x^2} \).

**Example:** Consider \( F(x,y,z) = x^2 + y^2 + z^2 - 2xy - 2xz - 2yz \). Consider the level set \( F(x,y,z) = 0 \). Compute \( \frac{\partial x}{\partial y} \) at point \((x,y,z) = (x_0, 1, 1)\) where \( x_0 \neq 0 \). (Comment: This means that \( x \) is regarded as a function of \( y \) and \( z \). If \( x,y \) are chosen to be independent, then \( \partial x/\partial y = 0 \) but here the problem obviously means they are dependent.)

**Solution.** By the implicit differentiation:
\[ \frac{\partial x}{\partial y} = -\frac{F_y}{F_x} = -\frac{2y - 2x - 2z}{2x - 2y - 2z} = -\frac{y - x - z}{x - y - z} = \ldots \]

We need to figure out \( x_0 \). \( \square \)

**Example:** In physics, for a gas system, \( P,V,T \) are the pressure, the volume and temperature respectively and they are related to each other through a constraint \( F(P,V,T) = 0 \) (for ideal gas, this is \( PV - nRT = 0 \) where \( n \) and \( R \) are constants). Show that
\[ \frac{\partial P \partial V \partial T}{\partial V \partial T \partial P} = -1. \]

You should understand what \( \frac{\partial P}{\partial V} \) means. This means \( P \) is regarded as a function of \( V \) and \( T \) and we are taking derivative on \( V \) by fixing \( T \).
Proof.

\[
\begin{align*}
\frac{\partial P}{\partial V} &= -\frac{F_V}{F_P} \\
\frac{\partial V}{\partial T} &= -\frac{F_T}{F_V} \\
\frac{\partial T}{\partial P} &= -\frac{F_P}{F_T}.
\end{align*}
\]

The product is $-1$.

You can verify it directly for the ideal gas case: $P = nRT/V$. Hence, $\partial P/\partial V = -nRT/V^2$. Similarly you can compute others. The product is $-1$. \qed

**Exercise:** In statistical physics, usually a system of $N$ particles is studied. The entropy $S$ depends on the energy $E$, the volume $V$ and $N$. Also, one can regard $S$ as the independent variable and $E$ is the dependent variable. Hence, $E$ then depends on $S, V, N$. Now, fixing $N$, show that

\[
\frac{(\partial S}{\partial V})_{N,E} = \frac{\partial S}{\partial E})_{N,V} = -(\frac{\partial E}{\partial V})_{N,S}.
\]

You cannot cancel the two $\partial S$ and get $\partial E/\partial V$. As you see, we have a negative sign. This is because the two $\partial S$'s are different changes.

**13.8 (Part 2): The gradient vector**

Recall the definition of gradient vector for $f(x, y)$, where $(x, y)$ are the Cartesian coordinates, is given by

\[
\nabla f(x, y) = \langle f_x(x, y), f_y(x, y) \rangle.
\]

We have similar expression for a function depends on $n$ variables.

The **chain rule** for $w = f(x(t), y(t), z(t))$ can be written as

\[
\frac{d}{dt} f(r(t)) = \nabla f(r(t)) \cdot r'(t) = (\nabla f \cdot u)v = (D_u f)v = \frac{dw}{ds} \frac{ds}{dt}.
\]

The first term measures how fast the functions changes with respect to distance while the second term measures how fast the the distance changes as time increases. Hence, the time rate is equal to the directional derivative times speed.

**Example:** Suppose $z = f(x, y) = 1 - x^2 - y^2$ describes the shape of a hill. You are at $(1, 0, 0)$. What is your rising rate (rise over time) if you head in the direction specified by $\vec{v} = \langle 2, 1 \rangle$ with speed $v = 1.5$? (Compare with the example in Lecture 7.)
Solution.
\[ \frac{df}{dt} = \nabla f \cdot \vec{r}' = v \cdot Du_f. \]
The speed is \( v = 1.5 \) while \( Du_f = \nabla f \cdot u \), which we computed in Lecture 10. □

Significance of the gradient vector

Recall \( Du_f = \nabla f \cdot u = |\nabla f| \cos \theta \). When \( \theta = 0 \), the change of rate is the fastest. The magnitude indicates the rate of change with respect distance.

1. The gradient points the fastest increasing direction. The magnitude indicates how fast the function values changes on distance.

When \( \theta = \pi/2 \), the rate of change is zero. In which direction is the changing rate zero? Along the level set! Hence, \( \nabla f \) is perpendicular with the level set. Actually, suppose \( \vec{r}(t) \) is a curve in the level set \( f = k \). By the chain rule \( 0 = \frac{d}{dt} f(\vec{r}) = \nabla f(\vec{r}) \cdot \vec{r}'(t) \). \( \vec{r}'(t) \) is tangent to the level set and hence \( \nabla f \) is normal to the level set.

2. The gradient vector is a normal vector of the level set.

Example: Consider the level set \( F(x,y) = x^3 + y^3 - 3xy = 0 \). At the two point \((\sqrt{4}, \sqrt{2})\) and \((0,0)\), \( F_y = 0 \), we can’t solve \( y \) in terms of \( x \). What happens at these two points? At the former, the tangent line is vertical. At the latter, there are two branches. By the geometric meaning, at \((0,0)\), \( \nabla F = 0 \).

Tangent planes revisited

Previously, we regarded the tangent plane as the linear approximation. Since any surface is the level set of some function. If we can find such a function, then, by taking the gradient of this function, we can find a normal vector of the plane.

Example: Find a normal vector of the graph of \( z = f(x,y) \).

Solution. \( F = f(x,y) - z \). Then, the graph is the zero level set of \( F \). The normal vector is simply \( \nabla F = (f_x, f_y, -1) \). From here, we can compute the tangent plane directly at \((a,b,f(a,b))\):

\[ \nabla F(a,b,f(a,b)) \cdot (x-a,y-b,z-f(a,b)) = 0. \]

Previously, we have the tangent plane to be \( z - f(a,b) = f_x(a,b)(x-a) + f_y(a,b)(y-b) \). Hence, the normal vector at \((a,b,f(a,b))\) is \( \vec{n} = \langle f_x(a,b), f_y(a,b), -1 \rangle \). This agrees.
Example: Consider $F(x, y, z) = \ln(x^2 + y^2 + z^2)$. The level set $F(x, y, z) = \ln(3)$ is a surface in 3D space passing through (1,1,1). Find the tangent plane of this surface at (1,1,1).

Solution. A normal vector is $n = \nabla F(1,1,1)$. Hence, the tangent plane is

$$\nabla F(1,1,1) \cdot (x - 1, y - 1, z - 1) = 0.$$

Similar techniques can be applied to tangent lines of $f(x, y) = k$. 

\[\square\]