Calculus

Notes on 2023/06/15

shaozewxy

June 2023

1 1.7 Derivatives with multiple variables

The purpose of calculus is to replace non-linear mappings with **linears transformations in small localities.**

1.1 Linear approximation in one-dimension

Given $U \subseteq R$ an open subset, and $f : U \to \mathbb{R}$, f is **differentiable** at $a \in U$ with derivative

$$f'(a) = \lim_{h \to 0} \frac{1}{h} \left(f(a+h) - f(a) \right)$$

if such limit exists.

1.2 Partial derivatives in $\mathbb{R}^1$

Here we do the same thing as in one-dimensional case.

Given $U \subseteq \mathbb{R}^n$ an open subset and $f : U \to \mathbb{R}$, we denote the derivative of f

w.r.t the i-th variable

$$D_i f(a) = \lim_{h\to 0} \frac{1}{h} \left(f \begin{pmatrix} a_1 \\ \ldots \\ a_i + h \\ \ldots \end{pmatrix} - f \begin{pmatrix} a_1 \\ \ldots \\ a_i \\ \ldots \end{pmatrix} \right)$$

i.e. fixing all other variables as constant and changing only the i-th variable.

1.3 Partial derivatives in $\mathbb{R}^m$

Defined similarly as partial derivatives in $\mathbb{R}^1$:

$$\overrightarrow{D_i f(a)} = \lim_{h\to 0} \frac{1}{h} \left(\mathbf{f} \begin{pmatrix} a_1 \\ \ldots \\ a_i + h \\ \ldots \end{pmatrix} - \mathbf{f} \begin{pmatrix} a_1 \\ \ldots \\ a_i \\ \ldots \end{pmatrix} \right) = \begin{bmatrix} D_i f_1(a) \\ \ldots \\ D_i f_m(a) \end{bmatrix}$$

1.4 Derivatives in several variables

With partial derivatives, can develop derivatives in several variables, i.e. how a system changes when all its components change.

In the case of one variable, derivative is defined as

$$\lim_{h\to 0} \frac{f(a+h) - f(a)}{h}$$

However, this definition doesn't work in multiple variables as $\mathbf{h}$ is a vector and can't be used to do division.

Therefore we rewrite the definition as:

Given $f : U \to \mathbf{R}$, f is differenciable at a with $f' = m$ if and only if:

$$\lim_{h \to 0} \frac{1}{h}((f(a+h) - f(a)) - mh) = 0$$

Because this limit equals 0, we can further rewrite it to

$$\lim_{h \to 0} \frac{1}{|h|}((f(a+h) - f(a)) - mh) = 0$$

With this definition, it can be expanded to multiple variables.

For $f : U \to \mathbb{R}^m, U \subseteq \mathbb{R}^n$, the linear transformation is defined by the **Jacobian matrix**:

$$Jf(a) = \begin{bmatrix} D_1 f_1(a) & ... & D_n f_1(a) \\ ... & & ... \\ D_1 f_m(a) & ... & D_n f_m(a) \end{bmatrix}$$

However, it is possible for some f to have all partial derivatives but can calculate $\lim_{h \to 0}(...)$.

Therefore, we define the derivative of multiple variables as:

Given $f : U \to \mathbb{R}^m, U \subseteq \mathbb{R}^n$, U open, $a \in U$. If $\exists$ a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h \to 0} \frac{1}{|h|}((f(a+h) - f(a)) - L(h)) = 0$$

Then f differentiable at a, and such L is represented by $Jf(a)$.

Proof:

NTS that $L = Jf(a)$. By definition we know that L is represented by the matrix:

$$\begin{bmatrix} L(e_1)...L(e_n) \end{bmatrix}$$

Therefore, we only NTS $L(e_i) = D_i f(a)$:

Let h approach 0 from e_i:

$$\lim_{te_i \longrightarrow 0} \frac{1}{|te_i|}((f(a+te_i) - f(a)) - L(te_i)) = 0$$

Because $|e_i| = 1$, we know that $|te_i| = |t||e_i| = |t|$,

$$\begin{aligned} &\lim_{te_i \to 0} \frac{1}{|t|}((f(a+te_i) - f(a)) - L(te_i)) \\ =&\lim_{te_i \to 0} \frac{f(a+te_i) - f(a)}{|t|} - \frac{L(te_i)}{|t|} \end{aligned}$$

Because the limit goes to 0, we can then replace $|t|$ with t:

$$\begin{aligned} &= \lim_{te_i \to 0} \frac{f(a+te_i) - f(a)}{t} - \lim_{te_i \to 0} \frac{tL(e_i)}{t} \\ &= \lim_{te_i \to 0} \frac{f(a+te_i) - f(a)}{t} - L(e_i) \end{aligned}$$

Since $\lim_{te_i \to 0} \frac{f(a+te_i)-f(a)}{t}$ is by definition $D_i f(a)$, therefore

$$D_i f(a) = L(e_i)$$

i.e. $Jf(a) = L$

1.4.1 Example using Jacobian Matrix

Take $f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} xy \\ x^2 - y^2 \end{pmatrix}$, denote the increment $\mathbf{v} = \begin{bmatrix} h \\ k \end{bmatrix}$, the Jacobian matrix is

$$\mathbf{Jf} = \begin{bmatrix} y & x \\ 2x & -2y \end{bmatrix}$$

Using the definition, we have:

$$\lim_{\mathbf{v}\to 0} \frac{1}{|\mathbf{v}| = \sqrt{h^2+k^2}} \left(\begin{pmatrix} (x+h)(y+h) \\ (x+h)^2 - (y+h)^2 \end{pmatrix} - \begin{pmatrix} xy \\ x^2 - y^2 \end{pmatrix} - \begin{bmatrix} y & x \\ 2x & -2y \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} \right)$$

should equal to 0.

$$\begin{aligned} &= \frac{1}{\sqrt{h^2+k^2}} \left(\begin{bmatrix} xy + hk + xk + hy - xy - yh - xk \\ x^2 + h^2 + 2xh - y^2 - k^2 - 2yk - x^2 + y^2 - 2xh + 2yk \end{bmatrix} \right) \\ &= \frac{1}{\sqrt{h^2+k^2}} \begin{bmatrix} hk \\ h^2 - k^2 \end{bmatrix} \end{aligned}$$

With $|h|, |k| \leq \sqrt{h^2+k^2}$, we have

$$\begin{aligned} \frac{hk}{\sqrt{h^2+k^2}} &= \frac{h}{\sqrt{h^2+k^2}} \cdot k \\ &\leq \frac{|h|}{\sqrt{h^2+k^2}} \cdot |k| \\ &\leq 1 \cdot \sqrt{h^2+k^2} \end{aligned}$$

Since $\mathbf{v} = \begin{bmatrix} h \\ k \end{bmatrix} \to 0$, we know $\frac{hk}{\sqrt{h^2+k^2}} \to 0$.

Similarly,

$$\begin{aligned} \frac{h^2-k^2}{\sqrt{h^2+k^2}} &= \frac{h^2}{\sqrt{h^2+k^2}} - \frac{k^2}{\sqrt{h^2+k^2}} \\ &\leq \frac{|h^2|}{\sqrt{h^2+k^2}} + \frac{|k^2|}{\sqrt{h^2+k^2}} \\ &\leq |h| \frac{|h|}{\sqrt{h^2+k^2}} + |k| \frac{|k|}{\sqrt{h^2+k^2}} \\ &\leq |h| + |k| \to 0 \end{aligned}$$

Therefore the limit above goes to 0, i.e. $\mathbf{Jf}$ is the derivative we want.

1.5 Directional derivatives

Finding the direvative over s certain direction other than the elemental directions can be done with Jacobian matrix:

Given $U \subseteq \mathbb{R}^n$ open, $f : U \to \mathbb{R}^m$ differenciable, the directional derivative of f over direction $\mathbf{v}$ is:

$$\lim_{h\to 0} \frac{f(a+h\mathbf{v}) - f(a)}{h}$$

and to compute this,

$$\lim_{h\to 0} \frac{f(a+h\mathbf{v}) - f(a)}{h} = [\mathbf{Jf}(\mathbf{a})]\mathbf{v}$$

Proof:

From the definition, we know that

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{1}{|\mathbf{h}|}\left(f(a+h) - f(a) - Jf(a)h\mathbf{v}\right) = 0$$

Substitute $\mathbf{h}$ with $h\mathbf{v}$, we have

$$\lim_{h\to 0} \frac{1}{|h\mathbf{v}|}(f(a+h\mathbf{v}) - f(a) - Jf(a)h\mathbf{v}) = 0$$
$$\lim_{h\to 0} \frac{1}{h}(f(a+h\mathbf{v}) - f(a) - Jf(a)h\mathbf{v}) = 0\cdot|\mathbf{v}| = 0$$

Therefore we can say that $\lim_{h\to 0} \frac{1}{h}(f(a+h\mathbf{v}) - f(a)) = Jf(a)\mathbf{v}$

1.5.1 Example of directional derivative

Given $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} xy \\ x^2 - y^2 \end{pmatrix}$ at $\mathbf{a} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ with dirction $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, then we have:

$$Jf = \begin{bmatrix} y & x \\ 2x & -2y \end{bmatrix}, Jf(\mathbf{a}) = \begin{bmatrix} 1 & 1 \\ 2 & -2 \end{bmatrix}, Jf(\mathbf{a})\mathbf{v} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Given $f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = xy\sin z$ over direction $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ at $\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ \pi/2 \end{bmatrix}$, the Jacobian matrix gives

$$Jf(\mathbf{a})\mathbf{v} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = 3$$

Using the definition, we have

$$\begin{aligned}
\lim_{h\to 0} \frac{1}{|h|}\left(f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a})\right) &= \frac{1}{|h|}((1+h)(1+2h)\cos(h+\pi/2) - 1) \\
&= \frac{1}{|h|}((2h^2+3h+1)(\cos h \sin \pi/2 + \sin h \cos \pi/2) - 1) \\
&= \frac{1}{|h|}(2h^2\cos h + 3h\cos h + (\cos h - 1)) \\
&= 2|h| + 2\cos h + \frac{\cos h - 1}{|h|} \\
&= 0 + 3 + 0 = 3
\end{aligned}$$

1.6 Matrix derivatives

Sometimes matrix derivatives are easier to compute using

$$\lim_{\mathbf{h}\to 0} \frac{1}{|\mathbf{h}|}(f(a+h) - f(a) - L(\mathbf{h}))$$

than does using Jacobian matrix.

Given $S : \mathrm{Mat}(n,n) \to \mathrm{Mat}(n,n), S(A) = A^2$, WTS $DS(A)$ is a linear transformation defined by $DS(A)H = AH + HA$:

$$
\begin{aligned}
\lim_{H\to 0}\frac{1}{|H|}((A+H)^2-A^2-DS(A)H) &= \frac{1}{|H|}((A+H)^2-A^2-AH-HA)\\
&= \frac{1}{|H|}(H^2+AH+HA-AH-HA)\\
&= |H| \to 0
\end{aligned}
$$

Given $S : \mathrm{Mat}(n,n) \to \mathrm{Mat}(n,n), S(A) = A^{-1}$, WTS $DS(A)H = -A^{-1}HA^{-1}$:

The definition gives:

$$
\lim_{H\to 0}\frac{1}{|H|}((A+H)^{-1}-A^{-1}+A^{-1}HA^{-1})
$$

Use the fact that if $|B| < 1 \Rightarrow I + B + B^2 + ... = (I-B)^{-1}$, we know as $|H| \to 0, |-A^{-1}H| < 1$, therefore,

$$
\begin{aligned}
\lim_{H\to 0}(A+H)^{-1} &= (A(I-(-A^{-1}H)))^{-1}\\
&= (I-(-A^{-1}H))^{-1}A^{-1}\\
&= (I+(-A^{-1}H)+(-A^{-1}H)^2+...)A^{-1}\\
&= A^{-1}-A^{-1}HA^{-1}+((-A^{-1}H)^2+...)A^{-1}
\end{aligned}
$$

The definition the becomes $\lim_{H\to 0}\frac{1}{|H|}((-A^{-1}H)^2+...)A^{-1}$, because all terms have $|H|$ more than 2 degrees, this goes to 0.