Lecture 5: Gradient, directional derivative and Taylor approximation

Learning goals:

1. What is a gradient?
2. What is the geometric interpretation of a gradient?
3. What is a directional derivative?
4. How does the Taylor approximation generalize to the multivariable case?

Where to find the material?
Corral 2.4 (does not contain Taylor)
Guichard et friends 14.5 (does not contain Taylor)
Active Calculus 10.6 (does not contain Taylor)
Adams-Essex 13.7, 13.9

Gradient

- Let $f\colon D \subset \mathbb{R}^n \to \mathbb{R}$, $n \geq 2$, be the derivative of the point $\mathbf{x} \in D$.

Definition

The **gradient** of the function f at $\mathbf{x}$ is the vector

$$\nabla f(\mathbf{x}) = \operatorname{grad} f(\mathbf{x}) = \left(\frac{\partial}{\partial x_1} f(\mathbf{x}), \frac{\partial}{\partial x_2} f(\mathbf{x}), \ldots, \frac{\partial}{\partial x_n} f(\mathbf{x})\right) \in \mathbb{R}^n.$$

- The gradient tells us the direction of the fastest growth of the function f. (Why? We will discuss this soon.).
- If f is differentiable at all points of D, one can define a vector-valued function $\nabla f\colon D \to \mathbb{R}^n$.
- In the case $n = 3$ one can write

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}.$$

- In the case $n = 2$, the third term is dropped.
- The gradient is a special case of the Jacobian matrix at $m = 1$.

Example

What is the gradient of $f(x, y) = x^2 + y^2$?

- $\nabla f(x) = (2x, 2y) = 2x\mathbf{i} + 2y\mathbf{j}$.
- What does this vector field look like? (draw a sketch)
- How is the vector field positioned with respect to the level curves of the function?

Gradient and level curves

Theorem

Let $D \subset \mathbb{R}^2$, $(a, b) \in D$ and $f : D \to \mathbb{R}$ be the differentiable at the point (a, b) such that $\nabla f(a, b) \neq \mathbf{0}$. Then $\nabla f(a, b)$ is perpendicular to the level curve (more precisely: to the tangent of the level curve) of f that goes through the point (a, b).

IMPORTANT!

Proof

Let $I = [-1, 1]$ and $\mathbf{r}(t)\colon I \to \mathbb{R}^2$, $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ curve such that

- $f(x(t), y(t)) = f(a, b) =$ constant for all $t \in I$ (i.e. $\mathbf{r}$ is a level curve)
- $\mathbf{r}(0) = (a, b)$

Consider now one variable function $f \circ \mathbf{r}(t) = f(x(t), y(t))$. This is a constant function.

Chain rule gives (since the derivative of the constant function is zero)

$$\begin{aligned} 0 &= \frac{d(f \circ \mathbf{r})}{dt}(t) = \partial_1 f\big(x(t), y(t)\big)x'(t) + \partial_2 f\big(x(t), y(t)\big)y'(t) \\ &= \nabla f(x(t), y(t)) \cdot \mathbf{r}'(t) \end{aligned}$$

In particular, at the point $t = 0$ this means that

$$\nabla f(a, b) \cdot \mathbf{r}'(0) = 0,$$

that is, the vector $\nabla f(a, b)$ and the tangent vector $\mathbf{r}'(0)$ of the curve $\mathbf{r}$ are perpendicular. $\square$

Directional derivative

- Partial derivatives give the growth rate of the function to the direction of the coordinate axes.
- For other directions, the growth rate is given by the *directional derivative* (if the limit exists)

$$D_{\mathbf{u}}f(a,b) = \lim_{h\to 0} \frac{f(a+hu_1, b+hu_2) - f(a,b)}{h}.$$

where $\mathbf{u} = (u_1, u_2)$ is the unit vector that gives the desired direction.

Theorem

Let $f\colon D \to \mathbb{R}$, $(a,b) \in D \subset \mathbb{R}^2$ and $\mathbf{u} = (u_1, u_2)$ such that $\|\mathbf{u}\|^2 = u_1^2 + u_2^2 = 1$. The directional derivative of the function f to the direction $\mathbf{u}$ is obtained from the formula

$$D_{\mathbf{u}}f(a,b) = \mathbf{u} \cdot \nabla f(a,b).$$

Can be proved by the chain rule

The growth rate of the function

- The directional derivatives (including partial derivatives) gives the growth rate to the specific direction.
- What direction gives the fastest growth?
- The definition of the dot product gives

$$D_{\mathbf{u}} f(a, b) = \mathbf{u} \cdot \nabla f(a, b) = \|\nabla f(a, b)\| \cos(\theta),$$

  where θ is the angle between the vectors $\mathbf{u}$ and $\nabla f(a, b)$.
- The highest value is thus obtained when $\theta = 0$ i.e. exactly to the direction of the gradient.
- And the growth rate to the direction of the gradient is exactly $\|\nabla f(a, b)\|$.

Example of a directional derivative

- Let $f(x, y) = y^4 + 2xy^3 + x^2y^2$. Find $D_{\mathbf{u}}f(0, 1)$ when $\mathbf{u}$ is the vector to the same direction than
  (a) $\mathbf{i} + 2\mathbf{j}$
  (b) $-2\mathbf{i} + \mathbf{j}$
  (c) $3\mathbf{i}$
  (d) $\mathbf{i} + \mathbf{j}$
- **Solution:** First calculate the gradient

$$\nabla f(x, y) = (2y^3 + 2xy^2)\mathbf{i} + (4y^3 + 6xy^2 + 2x^2y)\mathbf{j},$$

$$\nabla f(0, 1) = 2\mathbf{i} + 4\mathbf{j}.$$

(a) $\|\mathbf{i} + 2\mathbf{j}\| = \sqrt{5}$, a unit vector is needed: and hence $\mathbf{u} = (\mathbf{i} + 2\mathbf{j})/\sqrt{5}$. Thus

$$D_{\mathbf{u}}f(0, 1) = \frac{1}{\sqrt{5}}(\mathbf{i} + 2\mathbf{j}) \cdot (2\mathbf{i} + 4\mathbf{j}) = \frac{2 + 8}{\sqrt{5}} = 2\sqrt{5}.$$

Note that $\mathbf{u}$ and $\nabla f(0, 1)$ are parallel.

Example of a directional derivative (continues)

(b) $||\mathbf{j} - 2\mathbf{i}|| = \sqrt{5}$ and so $\mathbf{u} = (\mathbf{j} - 2\mathbf{i})/\sqrt{5}$. Thus

$$D_{\mathbf{u}}f(0,1) = \frac{1}{\sqrt{5}}(\mathbf{j} - 2\mathbf{i}) \cdot (2\mathbf{i} + 4\mathbf{j}) = \frac{-4+4}{\sqrt{5}} = 0.$$

The vectors $\mathbf{u}$ and $\nabla f(0,1)$ are therefore perpendicular.

(c) $||3\mathbf{i}|| = 3$ and therefore $\mathbf{u} = \mathbf{i}$. We get $D_{\mathbf{u}}f(0,1) = \mathbf{i} \cdot (2\mathbf{i} + 4\mathbf{j}) = 2$. This is equal to $\partial_1 f(0,1)$.

(d) $||\mathbf{i} + \mathbf{j}|| = \sqrt{2}$ and thus $\mathbf{u} = (\mathbf{i} + \mathbf{j})/\sqrt{2}$. Thus

$$D_{\mathbf{u}}f(0,1) = \frac{1}{\sqrt{2}}(\mathbf{i} + \mathbf{j}) \cdot (2\mathbf{i} + 4\mathbf{j}) = \frac{2+4}{\sqrt{2}} = 3\sqrt{2}.$$

Note that $3\sqrt{2} \approx 4.243 < 2\sqrt{5} \approx 4.472$.

Taylor polynomials

- Recall: For a single variable $m+1$ times the continuously derivable function $f\colon I \to \mathbb{R}$ can be approximated by the formula

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(m)}(a)}{m!}(x-a)^m.$$

or

$$f(a+h) \approx f(a) + f'(a)h + \frac{f''(a)}{2!}h^2 + \ldots + \frac{f^{(m)}(a)}{m!}h^m.$$

when $a, x \in I$ and $h = x - a$.

- Main idea behind finding the polynomial was: The values of derivatives have to be same to the m order derivatives
- This idea generalizes to the multivariate case.

- Let's calculate the second degree Taylor polynomial for two variable function $f = f(x, y)$ that has continuous 2nd order partial derivatives at the point (a, b).
- Second degree two variable polynomial is of the form

$$P_2(x, y) = a_0 + a_1(x-a) + a_2(y-b) + a_{1,1}(x-a)^2 + a_{1,2}(x-a)(y-b)$$

$$+\mathrm{a}_{2,2}(y-b)^2$$

- Let's match the values of derivatives:
- 0th order i.e. $f(a, b) = P_2(a, b)$
- 1st order i.e. $\partial_1 f(a, b) = \partial_1 P_2(a, b)$ and $\partial_2 f(a, b) = \partial_2 P_2(a, b)$
- 2nd order i.e. $\partial_{11} f(a, b) = \partial_{11} P_2(a, b)$, $\partial_{12} f(a, b) = \partial_{12} P_2(a, b)$ and $\partial_{22} f(a, b) = \partial_{22} P_2(a, b)$
- We obtain

$$P_2(x, y) = f(a, b) + \partial_1 f(a, b)(x-a) + \partial_2 f(a, b)(y-b) +$$

$$\frac{1}{2}\left(\partial_{11} f(a, b)(x-a)^2 + 2\partial_{12} f(a, b)(x-a)(y-b) + \partial_{22} f(a, b)(y-b)^2\right)$$

1st degree Taylor polynomial

- Similarly as above we get for two variable function $f = f(x, y)$ that has continuous 1st order partial derivatives at the point (a, b).

$$P_1(x, y) = f(a, b) + \partial_1 f(a, b)(x - a) + \partial_2 f(a, b)(y - b)$$

- This is familiar linear approximation

3rd degree Taylor polynomial

$$P_3(x,y) = f(a,b) + \partial_1 f(a,b)(x-a) + \partial_2 f(a,b)(y-b) +$$
$$+\frac{1}{2!}(\partial_{11} f(a,b)(x-a)^2 + 2\partial_{12} f(a,b)(x-a)(y-b) + \partial_{22}(y-b)^2)$$

$$+\frac{1}{3!}(\partial_{111} f(a,b)(x-a)^3 + 3\partial_{112} f(a,b)(x-a)^2(y-b)$$
$$+ 3\partial_{122} f(a,b)(x-a)(y-b)^2 + \partial_{222} f(a,b)(y-b)^3$$

Use of Taylor polynomials

- On this course we only need 2nd degree Taylor polynomial
- Taylor polynomials are used to approximate the function
- Taylor polynomials and their error approximations are used in numerical calculus and for example in differential geometry
- In practice very often, the Taylor polynomial for multivariable function can be obtained using single variable Taylor polynomials

Example of using Taylor polynomial

- Find the 2nd order Taylor approximation for the function $f(x, y) = \sqrt{x^2 + y^3}$ at the point $(1, 2)$.
- First $f(1, 2) = 3$,

$$\partial_1 f(x, y) = \frac{x}{\sqrt{x^2 + y^3}}, \quad \partial_2 f(x, y) = \frac{3y^2}{2\sqrt{x^2 + y^3}},$$

  that is, $\partial_1 f(1, 2) = 1/3$ and $\partial_2 f(1, 2) = 2$.
  Further

$$\partial_{11} f(x, y) = \frac{y^3}{(x^2 + y^3)^{3/2}}, \quad \partial_{11} f(1, 2) = \frac{8}{27},$$

Example of using Taylor polynomial

$$\partial_{12} f(x,y) = \frac{-3xy^2}{2(x^2+y^3)^{3/2}}, \quad \partial_{12} f(1,2) = -\frac{2}{9},$$

$$\partial_{22} f(x,y) = \frac{12x^2y+3y^4}{4(x^2+y^3)^{3/2}}, \quad \partial_{22} f(1,2) = \frac{2}{3}.$$

Thus

$$f(x,y) \approx 3 + \frac{1}{3}(x-1) + 2(y-2) + \frac{1}{2}\Big(\frac{8}{27}(x-1)^2 + 2\cdot(-\frac{2}{9})(x-1)(y-2) + \frac{2}{3}(y-2)^2\Big).$$

Another example

Find the 2nd degree Taylor polynomial for $f(x,y) = ye^{x-2y}$ at the point $(0,0)$.